

Evidence based Medicine -BCQ'S.

MRC EM-SUCCESS.



An FY2 in ED is collecting retrospective data for an audit on headache. He records the number of patients with headaches seen per day and for each patient who presented to ED with a headache he collects data on the length of time waiting to see the triage nurse, the pain score at triage (Faces Pain Rating Scale) and pain relief given in the ED (e.g. simple non-opioid analgesia, opioid analgesia and nil).

The number of patients with headaches seen per day is an example of which of the following types of data:

Categorical

Nominal

Ordinal

Discrete

An FY2 in ED is collecting retrospective data for an audit on headache. He records the number of patients with headaches seen per day and for each patient who presented to ED with a headache he collects data on the length of time waiting to see the triage nurse, the pain score at triage (Faces Pain Rating Scale) and pain relief given in the ED (e.g. simple non-opioid analgesia, opioid analgesia and nil).

The number of patients with headaches seen per day is an example of which of the following types of data:

Categorical

Nominal

Ordinal

Discrete

Continuous

 QUESTION 2

The interquartile range (IQ) is often displayed using which of the following:

Box and whisker plot

Funnel plot

Forest plot

Leaf and stem diagram

Scatter plot



See Answer



## QUESTION 2

The interquartile range (IQ) is often displayed using which of the following:

- |                                     |                       |     |
|-------------------------------------|-----------------------|-----|
| <input checked="" type="checkbox"/> | Box and whisker plot  | 77% |
| <input type="checkbox"/>            | Funnel plot           | 4%  |
| <input type="checkbox"/>            | Forest plot           | 7%  |
| <input type="checkbox"/>            | Leaf and stem diagram | 3%  |
| <input type="checkbox"/>            | Scatter plot          | 9%  |

## ANSWER

A box plot or box and whisker plot is used to display the interquartile range.





An FY2 in ED is collecting retrospective data for an audit on headache. He records the number of patients with headaches seen per day and for each patient who presented to ED with a headache he collects data on the length of time waiting to see the triage nurse, the pain score at triage (Faces Pain Rating Scale) and pain relief given in the ED (e.g. simple non-opioid analgesia, opioid analgesia and nil).

The type of analgesia given is an example of which of the following types of data:

Quantitative

Nominal

Ordinal

Discrete

Continuous

An FY2 in ED is collecting retrospective data for an audit on headache. He records the number of patients with headaches seen per day and for each patient who presented to ED with a headache he collects data on the length of time waiting to see the triage nurse, the pain score at triage (Faces Pain Rating Scale) and pain relief given in the ED (e.g. simple non-opioid analgesia, opioid analgesia and nil).

The type of analgesia given is an example of which of the following types of data:

Quantitative

Nominal

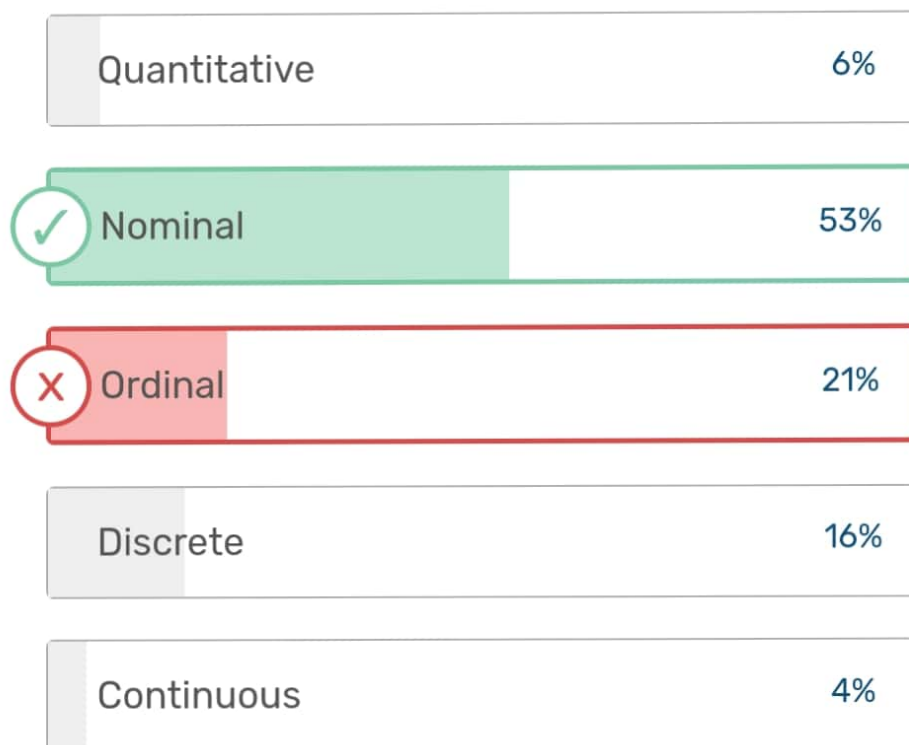
Ordinal

Discrete

Continuous



headaches seen per day and for each patient who presented to ED with a headache he collects data on the length of time waiting to see the triage nurse, the pain score at triage (Faces Pain Rating Scale) and pain relief given in the ED (e.g. simple non-opioid analgesia, opioid analgesia and nil). The type of analgesia given is an example of which of the following types of data:



## ANSWER

This is an example of nominal categorical data, where there is no order to the categorisation.





## QUESTION 4

Regarding bias, which of the following statements is CORRECT:

Bias is only introduced during the study design and analysis phases.

Bias is only important if the results of a study show an association between two variables.

Randomisation eliminates bias from a study.

Increasing the sample size reduces bias.


Bias leads to the systematic difference between the results from a study and the true states of affair.


Regarding bias, which of the following statements is CORRECT:

Bias is only introduced during the study design and analysis phases. 8%

Bias is only important if the results of a study show an association between two variables. 5%

Randomisation eliminates bias from a study. 19%

 Increasing the sample size reduces bias. 18%

 Bias leads to the systematic difference between the results from a study and the true states of affair. 51%


## ANSWER


Bias is the term used to describe an error at any stage of the study that was not due to chance. Bias leads to the systematic difference between the results from a study and the true states of affair. Bias may be introduced at all stages of the research process, from study design, through to

study design and analysis phases.

Bias is only important if the results of a study show an association between two variables. 5%

Randomisation eliminates bias from a study. 19%

 Increasing the sample size reduces bias. 18%

 Bias leads to the systematic difference between the results from a study and the true states of affair. 51%

## ANSWER

Bias is the term used to describe an error at any stage of the study that was not due to chance. Bias leads to the systematic difference between the results from a study and the true states of affair. Bias may be introduced at all stages of the research process, from study design, through to analysis and publication. Bias can create a spurious association or mask a real association.

Good research design can reduce the effect of bias (e.g. blinding, randomisation) but they cannot eliminate it completely. Increasing the sample size does not reduce bias.



Increasing the sample size reduces bias.

18%



Bias leads to the systematic difference between the results from a study and the true states of affair.

51%

## ANSWER

Bias is the term used to describe an error at any stage of the study that was not due to chance. Bias leads to the systematic difference between the results from a study and the true states of affair. Bias may be introduced at all stages of the research process, from study design, through to analysis and publication. Bias can create a spurious association or mask a real association.

Good research design can reduce the effect of bias (e.g. blinding, randomisation) but they cannot eliminate it completely. Increasing the sample size does not reduce bias.



Save

End Session





## QUESTION 5

A new blood test is being developed to diagnose DVT. 1000 people presenting to ED with suspected DVT undergo the new blood test and the gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the negative predictive value of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

A new blood test is being developed to diagnose DVT. 1000 people presenting to ED with suspected DVT undergo the new blood test and the gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the negative predictive value of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

86%

95%

98%

gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the negative predictive value of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a = 75  |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

86%

95%

98%

99%

the negative predictive value of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a = 75  |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

6%

86%

16%

X

95%

6%

98%

8%

✓

99%

64%

## ANSWER

- Negative predictive value (NPV) =  $d/(c+d) = 798/800 = 0.99 = 99\%$ 
  - This means there is a 99% chance, if the test is negative, that the patient does not have a DVT.





A new test is being developed to diagnose chlamydia. 1000 people aged 15 – 35 years attending a GUM clinic undergo the new test and the current gold standard nucleic acid amplification test (NAAT) to confirm the diagnosis. Of the 1000 people, 250 are diagnosed with chlamydia. Of the patients diagnosed with chlamydia, 240 test positive with the new diagnostic test and of the patients not diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the specificity of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |

aged 15-25 years attending a sexual health clinic undergo the new test and the current gold standard nucleic acid amplification test (NAAT) to confirm the diagnosis. Of the 1000 people, 250 are diagnosed with chlamydia. Of the patients diagnosed with chlamydia, 240 test positive with the new diagnostic test and of the patients not diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the specificity of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |

62%

75%

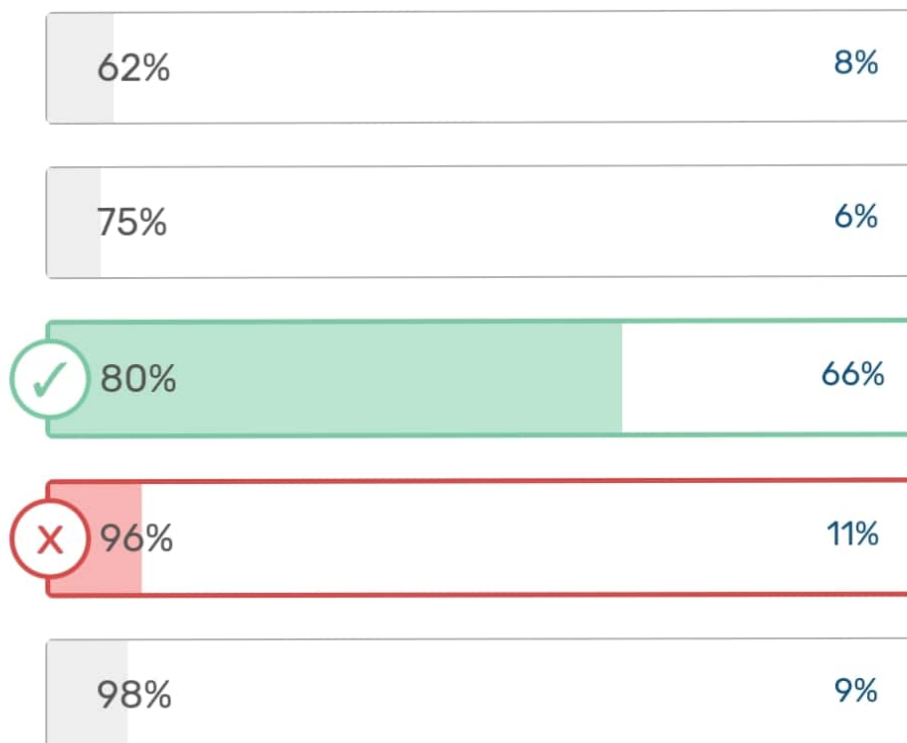
80%

96%

98%

diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the specificity of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |



## ANSWER

- Specificity =  $d/(b+d) = 600/750 = 0.8 = 80\%$ 
  - This means if the patient does not have chlamydia, there is an 80% chance of the test being negative. The test will have a 20% false positive result

|       |     |
|-------|-----|
| Total | 250 |
|-------|-----|



## ANSWER

- Specificity =  $d/(b+d) = 600/750 = 0.8 = 80\%$ 
  - This means if the patient does not have chlamydia, there is an 80% chance of the test being negative. The test will have a 20% false positive result.



Save

End Session



## QUESTION 7

A range of one standard deviation above and below the mean includes what approximate percentage of the sample values:

[See Answer](#)



## QUESTION 7

A range of one standard deviation above and below the mean includes what approximate percentage of the sample values:

56%

11%



68%

72%

72%

6%



78%

6%

84%

6%

## ANSWER

A range of one SD above and below the mean (+/- 1 SD) includes 68.2% of the sample values.



## QUESTION 8

The results from the study investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the sensitivity calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

$$a/(a+c)$$

$$d/(c+d)$$

$$a/(a+b)$$

$$(a+c)/c$$



## QUESTION 8

The results from the study investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the sensitivity calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

$$a/(a+c)$$

$$d/(c+d)$$

$$a/(a+b)$$

$$(a+c)/a$$

The results from the study investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the sensitivity calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

4%



$$a/(a+c)$$

79%

$$d/(c+d)$$

4%



$$a/(a+b)$$

11%

$$(a+c)/a$$

2%

## ANSWER

Sensitivity is the proportion of patients with the disease (true positives) who are correctly identified by the test as having the disease = true positive rate

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

4%



$$a/(a+c)$$

79%

$$d/(c+d)$$

4%



$$a/(a+b)$$

11%

$$(a+c)/a$$

2%

## ANSWER

Sensitivity is the proportion of patients with the disease (true positives) who are correctly identified by the test as having the disease = true positive rate.

$$\text{Sensitivity} = a/(a+c)$$





## QUESTION 9

Regarding epidemiology, which of the following statements is CORRECT:

Incidence is defined as the total number of cases of a disease per population at a given point in time.

In chronic disease prevalence is usually greater than incidence.

Prevalence is always greater than incidence.

Prevalence is defined as the total number of deaths attributed to a disease in a given time period.

Cross-sectional studies give a good estimate of incidence of a disease.

Regarding epidemiology, which of the following statements is CORRECT:

Incidence is defined as the total number of cases of a disease per population at a given point in time. 25%

✓ In chronic disease prevalence is usually greater than incidence. 49%

Prevalence is always greater than incidence. 10%

✗ Prevalence is defined as the total number of deaths attributed to a disease in a given time period. 7%

Cross-sectional studies give a good estimate of incidence of a disease. 9%

## ANSWER

The incidence of a disease is the number of new cases of a condition over a given time period, given as a percentage of the population. The point prevalence of a disease is the existing number of cases of a condition at a single point in time, given as a percentage of the population.

✓ usually greater than incidence.

Prevalence is always greater than incidence. 10%

✗ Prevalence is defined as the total number of deaths attributed to a disease in a given time period. 7%

Cross-sectional studies give a good estimate of incidence of a disease. 9%

## ANSWER

The incidence of a disease is the number of new cases of a condition over a given time period, given as a percentage of the population. The point prevalence of a disease is the existing number of cases of a condition at a single point in time, given as a percentage of the population. The mortality rate is a type of incidence rate defined as the total number of deaths attributed to a disease in a given time period in a population. With chronic diseases, the incidence will be lower than the prevalence. With short-term illnesses e.g. the common cold, the incidence will be greater than the prevalence. Cross-sectional studies can be used to estimate the prevalence of disease in a population.



## QUESTION 10

A range of three standard deviations above and below the mean includes what approximate percentage of the sample values:

[See Answer](#)



## QUESTION 10

A range of three standard deviations above and below the mean includes what approximate percentage of the sample values:

88%

4%

90%

3%



92%

6%

95%

14%



99%

73%

## ANSWER

A range of three SD above and below the mean (+/- 3 SD) includes 99.7% of the sample values.



QUESTION 11

A type I error occurs when:

The null hypothesis is accepted when it is true.

The null hypothesis is accepted when it is false.

The null hypothesis is rejected when it is true.

The alternative hypothesis is accepted when it is false.

The alternative hypothesis is rejected when it is true.





A type I error occurs when:

The null hypothesis is accepted when it is true. 7%

The null hypothesis is accepted when it is false. 19%

☒ The null hypothesis is rejected when it is true. 68%

The alternative hypothesis is accepted when it is false. 4%

The alternative hypothesis is rejected when it is true. 2%

## ANSWER

A type I error occurs when the null hypothesis is wrongly rejected when it is actually true and we conclude that there is a difference of effect when in reality there is none (a false positive result).



## QUESTION 12

You are reading about a potential new treatment for migraine as part of your continuing professional development. Which of the following values is needed in order to calculate the number needed to treat (NNT):

Absolute risk

Absolute risk reduction

Relative risk

Relative risk reduction

Number needed to harm





## QUESTION 12

You are reading about a potential new treatment for migraine as part of your continuing professional development. Which of the following values is needed in order to calculate the number needed to treat (NNT):

Absolute risk

13%



Absolute risk reduction

38%

Relative risk

38%

Relative risk reduction

6%

Number needed to harm

6%

## ANSWER

The **number needed to treat (NNT)** is the number of patients who need to be treated with the intervention, compared with the control, in order for one extra patient to experience a beneficial

Which of the following values is needed in order to calculate the number needed to treat (NNT):

Absolute risk

13%



Absolute risk reduction

38%

Relative risk

38%

Relative risk reduction

6%

Number needed to harm

6%

## ANSWER

The **number needed to treat (NNT)** is the number of patients who need to be treated with the intervention, compared with the control, in order for one extra patient to experience a beneficial effect.

It is the reciprocal of the absolute risk reduction and therefore gives us information about absolute benefit.

$$\text{NNT} = 1/\text{ARR}$$





A randomised controlled trial (RCT) is performed where 1000 men are treated with a lipid-lowering drug, Superstatin, and 1000 given a placebo. Two years later the number of patients who have had a myocardial infarction (MI) is recorded and shown below. What is the absolute risk reduction of treatment in preventing an MI:

|             | Myocardial infarction<br>Yes |
|-------------|------------------------------|
| Superstatin | 20                           |
| Placebo     | 40                           |
| Total       | 60                           |

0.5%

2%

4%

25%

50%

Superstatin, and 1000 given a placebo. Two years later the number of patients who have had a myocardial infarction (MI) is recorded and shown below. What is the absolute risk reduction of treatment in preventing an MI:

|             | Myocardial infarction<br>Yes |
|-------------|------------------------------|
| Superstatin | 20                           |
| Placebo     | 40                           |
| Total       | 60                           |

0.5%

2%

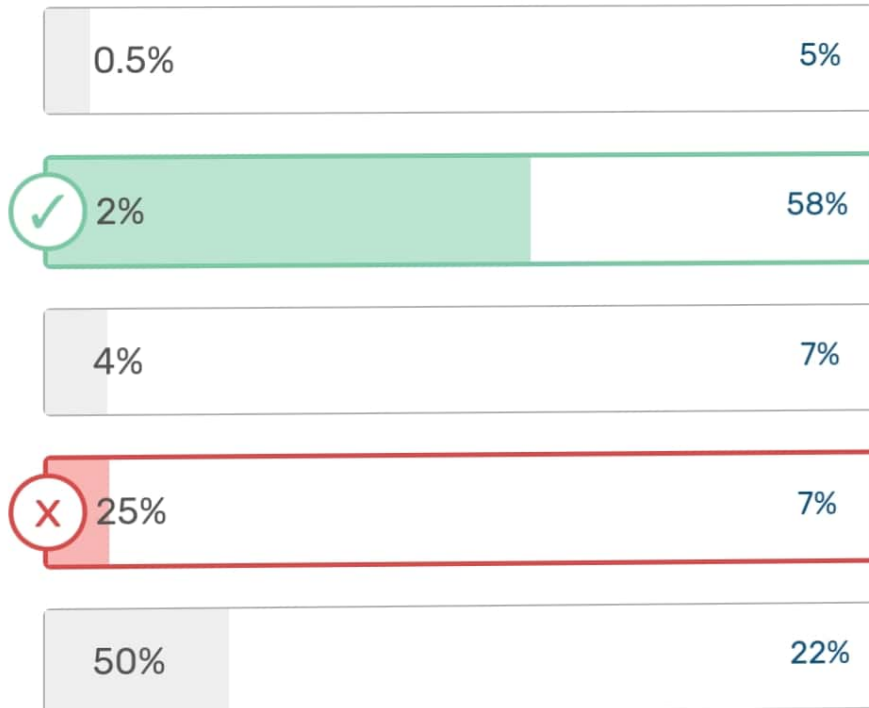
4%

25%

50%

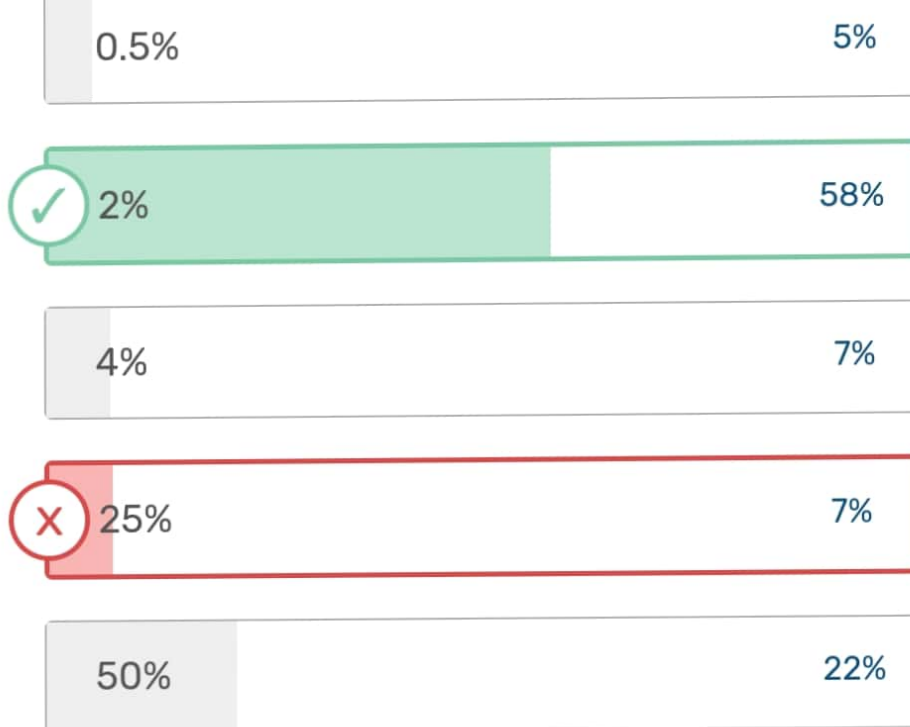


|             |    |
|-------------|----|
| Superstatin | 20 |
| Placebo     | 40 |
| Total       | 60 |



## ANSWER

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the



## ANSWER

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the absolute risk of having an MI by 2%.





## QUESTION 14

Regarding confidence intervals for the sample mean, which of the following statements is CORRECT:

The 95% confidence interval for the mean is the mean  $\pm 2.58 \times$  standard error of the mean (SEM).

A wider confidence interval shows a more precise estimate of the true population mean.

The width of the confidence interval is dependent on the sample size and the sample mean.


The width of the confidence interval narrows with decreasing sample size.

A 95% confidence interval is the range in which we can be 95% confident that the true population mean lies.


Regarding confidence intervals for the sample mean, which of the following statements is CORRECT:

The 95% confidence interval for the mean is the mean  $\pm 2.58 \times$  standard error of the mean (SEM). 10%

A wider confidence interval shows a more precise estimate of the true population mean. 6%

 The width of the confidence interval is dependent on the sample size and the sample mean. 19%

The width of the confidence interval narrows with decreasing sample size. 8%


 A 95% confidence interval is the range in which we can be 95% confident that the true population mean lies. 56%

## ANSWER


The 95% confidence interval for the mean is the mean  $\pm 1.96$  SEM. This is the range in which we can be 95% confident that the true population mean lies. The width of the confidence interval will decrease with a smaller SEM, which in turn

mean is the mean  $\pm 2.58 \times$  standard error of the mean (SEM).

A wider confidence interval shows a more precise estimate of the true population mean. 6%

 The width of the confidence interval is dependent on the sample size and the sample mean. 19%

The width of the confidence interval narrows with decreasing sample size. 8%

 A 95% confidence interval is the range in which we can be 95% confident that the true population mean lies. 56%

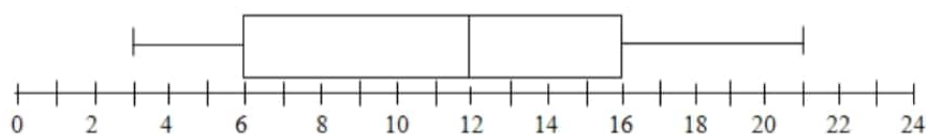
## ANSWER

The 95% confidence interval for the mean is the mean  $\pm 1.96$  SEM. This is the range in which we can be 95% confident that the true population mean lies. The width of the confidence interval will decrease with a smaller SEM, which in turn will decrease with increasing sample size or decreasing variability. A narrow confidence interval shows a more precise estimate.



## QUESTION 15

Regarding the following box and whisker plot, what is the range of the data:



6

12

16

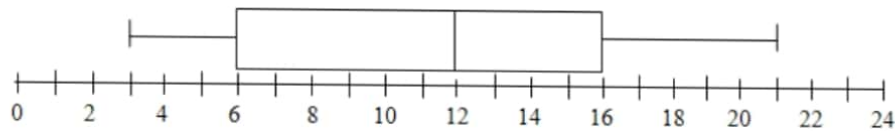
18

21





Regarding the following box and whisker plot, what is the range of the data:



6

3%

12

14%

X

16

6%

✓

18

68%

21

10%

## ANSWER

The range is represented by the whole of the box and whisker plot. The range is the difference between the maximum and minimum value, demarcated by the ends of the whiskers. The range =  $21 - 3 = 18$ .



## QUESTION 16

Regarding cross-sectional studies, which of the following statements is CORRECT:

They are useful for assessing causation of an variable on an outcome.

They are experimental studies.

They are particularly suitable for estimating point prevalence.

They are particularly suitable for estimating incidence of a disease.

They are useful for considering trends over time.



Regarding cross-sectional studies, which of the following statements is CORRECT:

They are useful for assessing causation of an variable on an outcome. 11%

They are experimental studies. 7%



They are particularly suitable for estimating point prevalence. 53%



They are particularly suitable for estimating incidence of a disease. 19%

They are useful for considering trends over time. 11%

## ANSWER

Cross-sectional studies aim to provide data about population health, normal ranges of biological parameters, and disease prevalence or severity by observing the entire population, or a representative subset, at a single point in time. Cross-sectional studies are relatively simple and

They are useful for assessing causation of an variable on an outcome. 11%

They are experimental studies. 7%

✓ They are particularly suitable for estimating point prevalence. 53%

✗ They are particularly suitable for estimating incidence of a disease. 19%

They are useful for considering trends over time. 11%

## ANSWER

Cross-sectional studies aim to provide data about population health, normal ranges of biological parameters, and disease prevalence or severity by observing the entire population, or a representative subset, at a single point in time. Cross-sectional studies are relatively simple and quick to perform and can be used to study multiple outcomes, but are subject to confounding and recall bias and are not suitable for studying rare diseases. Cross-sectional studies cannot be used to assess causation or to consider trends over time.



## QUESTION 17

Regarding the interpretation of risk of an outcome event when comparing a treatment and control group, how is the absolute risk reduction calculated:

Absolute risk in the control group -  
Absolute risk in the treatment group

Absolute risk in the control group /  
Absolute risk in the treatment group

1 - Absolute risk in the treatment  
group

Absolute risk in the treatment group  
/ Absolute risk in the control group

1 - Absolute risk in the control group



## QUESTION 17

Regarding the interpretation of risk of an outcome event when comparing a treatment and control group, how is the absolute risk reduction calculated:



Absolute risk in the control group - Absolute risk in the treatment group 58%



Absolute risk in the control group / Absolute risk in the treatment group 8%

1 - Absolute risk in the treatment group 9%

Absolute risk in the treatment group / Absolute risk in the control group 20%

1 - Absolute risk in the control group 4%

## ANSWER

The absolute risk reduction (ARR) is the difference between the absolute risk of outcome event in the control group (ARC) and the

the absolute risk reduction calculated:

✓ Absolute risk in the control group - 58%  
Absolute risk in the treatment group

✗ Absolute risk in the control group / 8%  
Absolute risk in the treatment group

1 - Absolute risk in the treatment 9%  
group

Absolute risk in the treatment group / 20%  
Absolute risk in the control group

1 - Absolute risk in the control group 4%

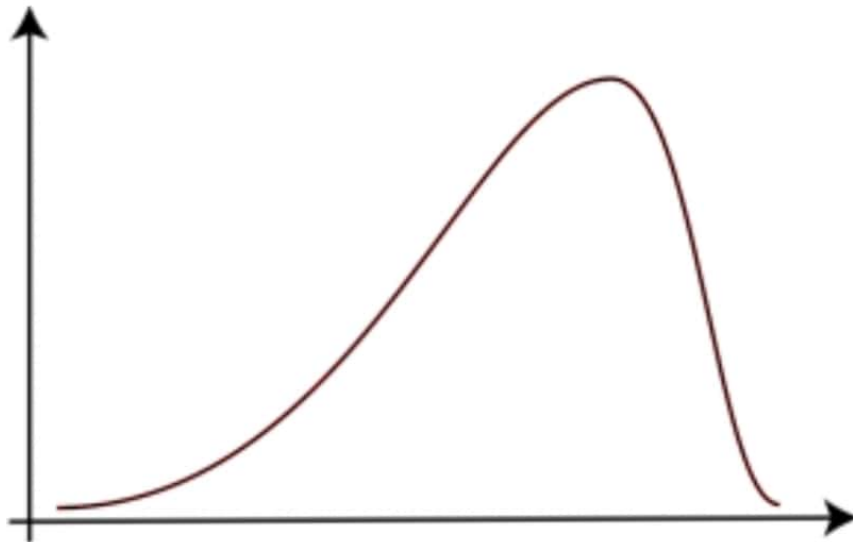
## ANSWER

The absolute risk reduction (ARR) is the difference between the absolute risk of outcome event in the control group (ARC) and the absolute risk of outcome event in the treatment group (ART).

$$ARR = ARC - ART$$



Regarding the following data distribution, which of the following statements is INCORRECT:



The distribution is negatively skewed.

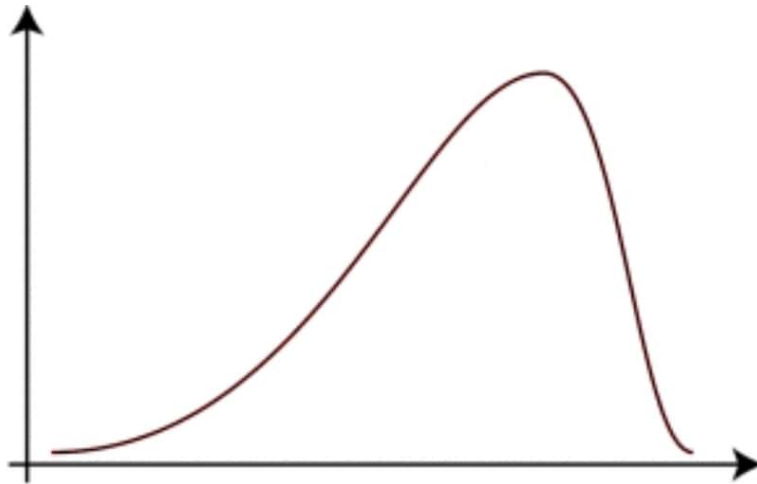
The distribution is unimodal.

The mean is less than the median.

The median is less than the mode.

The mean is greater than the mode.

statements is INCORRECT:



The distribution is negatively skewed. 14%

The distribution is unimodal. 6%

☒ The mean is less than the median. 13%

The median is less than the mode. 9%

☒ The mean is greater than the mode. 58%

## ANSWER

This unimodal distribution is negatively skewed. In a negative skew, the left tail is longer and the mass of distribution is concentrated on the right. The mean  $<$  median  $<$  mode.



QUESTION 19

Forest plots are typically used to display:

The median and interquartile range

The existence of publication bias in meta-analysis

The survival of a sample cohort

The effect size of constituent trials in a meta-analysis

The strength of correlation between two variables



See Answer




Forest plots are typically used to display:

The median and interquartile range 11%

The existence of publication bias in meta-analysis 19%

The survival of a sample cohort 10%

 The effect size of constituent trials in a meta-analysis 52%

The strength of correlation between two variables 9%

## ANSWER

Forest plots are usually found in meta-analysis and provide graphical representation of the strength of evidence of the constituent trials.



Save

End Session

Regarding the interpretation of risk of an outcome event when comparing a treatment and control group, how is the relative risk calculated:

Absolute risk in the control group -  
Absolute risk in the treatment group

Absolute risk in the control group /  
Absolute risk in the treatment group

Absolute risk in the treatment group  
- Absolute risk in the control group

Absolute risk in the treatment group  
/ Absolute risk in the control group

1 - Absolute risk in the control group



See Answer

an outcome event when comparing a treatment and control group, how is the relative risk calculated:

Absolute risk in the control group - 15%  
Absolute risk in the treatment group

Absolute risk in the control group / 13%  
Absolute risk in the treatment group

Absolute risk in the treatment group - 10%  
Absolute risk in the control group

✓ Absolute risk in the treatment group / 53%  
Absolute risk in the control group

1 - Absolute risk in the control group 10%

## ANSWER

The relative risk (RR) is the ratio of the absolute risk of outcome event in the treatment group (ART) compared to the absolute risk of outcome event in the control group (ARC).

$$RR = ART/ARC$$



 QUESTION 21

Which of the following risk ratios indicates no difference in risk between two groups:

[See Answer](#)

Which of the following risk ratios indicates no difference in risk between two groups:

|                                      |     |
|--------------------------------------|-----|
| -1                                   | 2%  |
| 0                                    | 34% |
| <input checked="" type="radio"/> 0.5 | 6%  |
| <input checked="" type="radio"/> 1   | 54% |
| $\infty$                             | 5%  |

## ANSWER

- A risk ratio of 1 indicates no difference in risk between groups.
- If the risk ratio of an event is  $> 1$ , the rate of that event is increased in the exposed group compared to the control group.
- If the risk ratio is  $< 1$ , the rate of that event is reduced in the exposed group compared to the control group.



Save

End Session



## QUESTION 22

If the null hypothesis is wrongly accepted when it is actually false, this is an example of:

A test with high power

A test with a high negative predictive value

A test with a type I error

A test with nonparametric data

A test with a type II error



See Answer



## QUESTION 22

If the null hypothesis is wrongly accepted when it is actually false, this is an example of:

A test with high power

2%

A test with a high negative predictive value

5%



A test with a type I error

26%

A test with nonparametric data

2%



A test with a type II error

66%

## ANSWER

A type II error occurs when the null hypothesis is wrongly accepted when it is actually false and we conclude that there is no evidence of a difference in effect when one really exists (a



If the null hypothesis is wrongly accepted when it is actually false, this is an example of:

A test with high power

2%

A test with a high negative predictive value

5%



A test with a type I error

26%

A test with nonparametric data

2%



A test with a type II error

66%

## ANSWER

A type II error occurs when the null hypothesis is wrongly accepted when it is actually false and we conclude that there is no evidence of a difference in effect when one really exists (a false negative result).



Save

End Session



### QUESTION 23

How is the standard error of the mean for a data sample calculated:

Mean  $\pm$  1.96 x Standard deviation

Standard deviation/ $\sqrt{\text{sample size}}$

Mean/ $\sqrt{\text{sample size}}$

Mean/ $\sqrt{\text{standard deviation}}$

Standard deviation x sample size



See Answer

Save

End Session



## QUESTION 23

How is the standard error of the mean for a data sample calculated:

Mean  $\pm$  1.96 x Standard deviation 20%



Standard deviation/ $\sqrt{\text{sample size}}$  46%



Mean/ $\sqrt{\text{sample size}}$  8%

Mean/ $\sqrt{\text{standard deviation}}$  22%

Standard deviation x sample size 4%

## ANSWER

The SEM for a data sample =  $\sigma/\sqrt{n}$ , where  $\sigma$  = standard deviation and  $n$  = sample size.



Regarding likelihood ratios, which of the following statements is INCORRECT:

Likelihood ratios, like predictive values, are affected by the prevalence of the disease in the population.

Positive likelihood ratio =  
 $\text{sensitivity} / (1 - \text{specificity})$

A test with a low negative likelihood ratio is good at increasing certainty about the absence of a disorder.

The positive likelihood ratio can be used to calculate the post-test odds of a disease.

Negative likelihood ratio =  $(1 - \text{sensitivity}) / \text{specificity}$



INCORRECT:

✓ Likelihood ratios, like predictive values, are affected by the prevalence of the disease in the population. 31%

Positive likelihood ratio =  $\text{sensitivity} / (1 - \text{specificity})$  10%

A test with a low negative likelihood ratio is good at increasing certainty about the absence of a disorder. 29%

✗ The positive likelihood ratio can be used to calculate the post-test odds of a disease. 18%


Negative likelihood ratio =  $(1 - \text{sensitivity}) / \text{specificity}$  13%

## ANSWER

A likelihood ratio is a measure of the diagnostic value of a test. Likelihood ratios show how many times more likely patients with a disease are to have a particular test result than patients without the disease. Likelihood ratios are more useful than predictive values because they are calculated from sensitivity and specificity and

Positive likelihood ratio = sensitivity/(1 - specificity) 10%

A test with a low negative likelihood ratio is good at increasing certainty about the absence of a disorder. 29%

 The positive likelihood ratio can be used to calculate the post-test odds of a disease. 18%

Negative likelihood ratio = (1 - sensitivity)/specificity 13%

## ANSWER

A likelihood ratio is a measure of the diagnostic value of a test. Likelihood ratios show how many times more likely patients with a disease are to have a particular test result than patients without the disease. Likelihood ratios are more useful than predictive values because they are calculated from sensitivity and specificity and therefore remain constant even when the prevalence of the disorder changes.



Save

End Session

 QUESTION 25

Which of the following is an example of a non-parametric test:

Mann-Whitney U test

Wilcoxon matched pairs test

Kruskal-Wallis test

Friedman's test

All of the above



See Answer

Save

End Session



## QUESTION 25

Which of the following is an example of a non-parametric test:

Mann-Whitney U test

6%

Wilcoxon matched pairs test

3%



Kruskal-Wallis test

4%

Friedman's test

6%



All of the above

81%

### ANSWER

All of the above answers are non-parametric tests.





## QUESTION 26

Which of the following terms describes the proportion of patients with a disease who are correctly identified by a test as having the disease:

Specificity

Sensitivity

Negative predictive value

Positive predictive value

Positive likelihood ratio



See Answer



## QUESTION 26

Which of the following terms describes the proportion of patients with a disease who are correctly identified by a test as having the disease:

Specificity

10%



Sensitivity

65%



Negative predictive value

2%

Positive predictive value

21%

Positive likelihood ratio

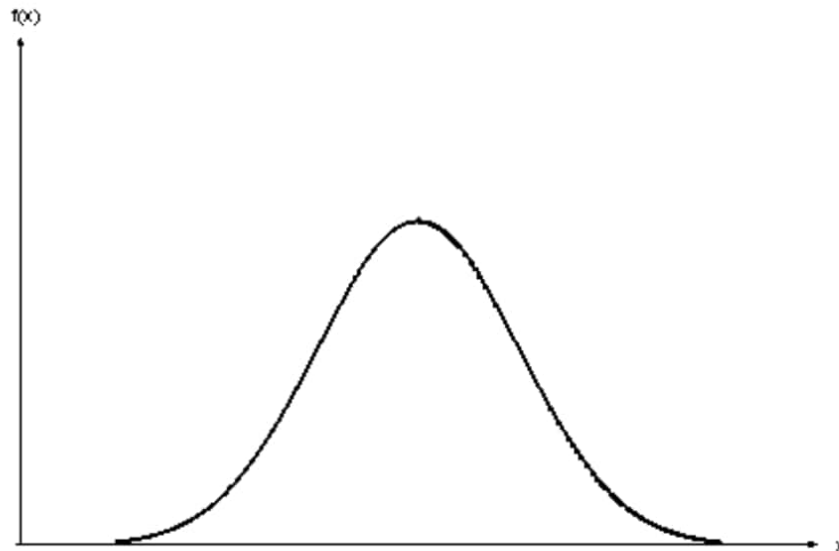
2%

## ANSWER

Sensitivity is the proportion of patients with the disease (true positives) who are correctly identified by the test as having the disease = true positive rate.



Regarding the following data distribution, which of the following statements is INCORRECT:



The distribution is unimodal.

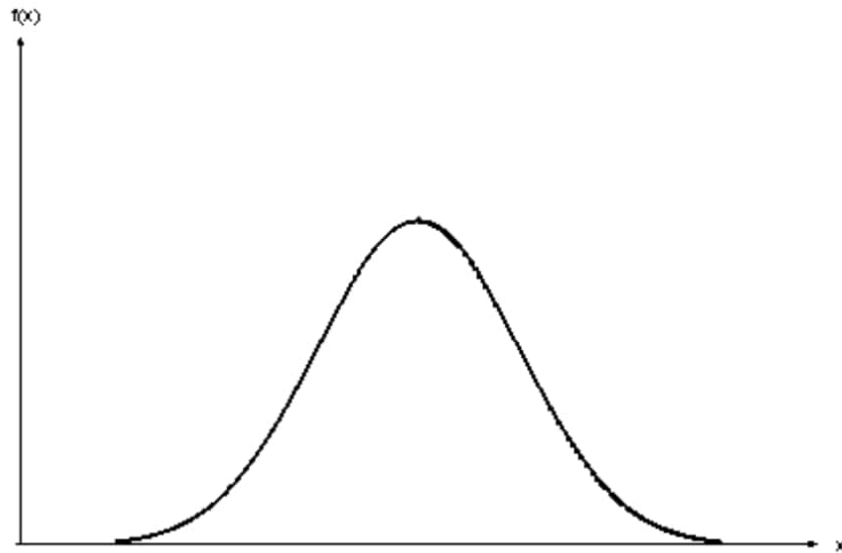
The mean = median.

The distribution will shift to the right if the mean is increased.

The distribution will become more peaked as the standard deviation is increased.

The distribution is a normal distribution.

statements is INCORRECT:



The distribution is unimodal.

The mean = median.

The distribution will shift to the right if the mean is increased.

The distribution will become more peaked as the standard deviation is increased.

The distribution is a normal distribution.



The distribution is unimodal.

7%

The mean = median.

7%

The distribution will shift to the right if the mean is increased.

19%



The distribution will become more peaked as the standard deviation is increased.

58%

The distribution is a normal distribution.

8%

## ANSWER

The normal distribution is:

- bell-shaped (unimodal)
- symmetrical about its mean (and the mean and median are equal)
- completely described by two parameters, the mean and the standard deviation of the population
- shifted to the right if the mean is increased and to the left if the mean is decreased (for a constant standard deviation)
- flattened as the standard deviation is increased but becomes more peaked as

The distribution will shift to the right if the mean is increased. 19%



The distribution will become more peaked as the standard deviation is increased. 58%

The distribution is a normal distribution. 8%

## ANSWER

The normal distribution is:

- bell-shaped (unimodal)
- symmetrical about its mean (and the mean and median are equal)
- completely described by two parameters, the mean and the standard deviation of the population
- shifted to the right if the mean is increased and to the left if the mean is decreased (for a constant standard deviation)
- flattened as the standard deviation is increased but becomes more peaked as the standard deviation is decreased (for a fixed mean)





QUESTION 28

A type II error occurs when:

The null hypothesis is accepted when it is true.

The null hypothesis is accepted when it is false.

The null hypothesis is rejected when it is true.

The alternative hypothesis is accepted when it is false.

The alternative hypothesis is rejected when it is true.





A type II error occurs when:

The null hypothesis is accepted when it is true. 4%

✓ The null hypothesis is accepted when it is false. 71%

✗ The null hypothesis is rejected when it is true. 17%

The alternative hypothesis is accepted when it is false. 6%

The alternative hypothesis is rejected when it is true. 3%

## ANSWER

A type II error occurs when the null hypothesis is wrongly accepted when it is actually false and we conclude that there is no evidence of a difference in effect when one really exists (a false negative result).



A new blood test is being developed to diagnose DVT. 1000 people presenting to ED with suspected DVT undergo the new blood test and the gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the positive likelihood ratio (LR+) for this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |

gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the positive likelihood ratio (LR+) for this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a = 75  |
| Negative test | c = 2   |
| Total         | 77      |

3.7

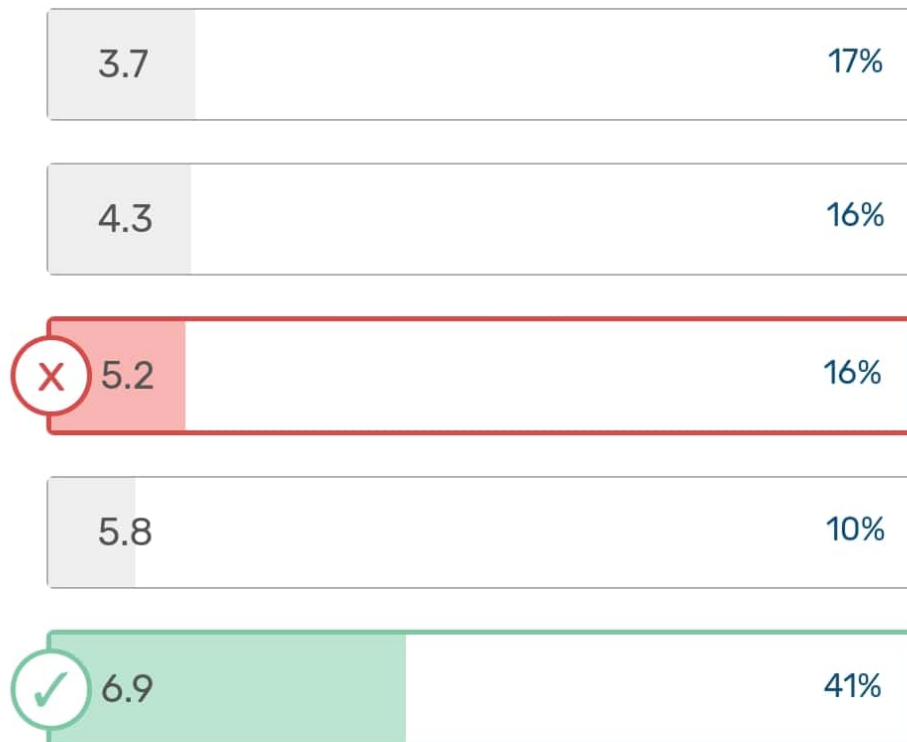
4.3

5.2

5.8

6.9

|               |       |
|---------------|-------|
| Negative test | c = 2 |
| Total         | 77    |



## ANSWER

- Sensitivity =  $a/(a+c) = 75/77 = 0.97 = 97\%$ 
  - This means that if the patient has a DVT, there is a 97% chance of the test being positive. The test will only have a 3% false negative rate.
- Specificity =  $d/(b+d) = 798/923 = 0.86 = 86\%$ 
  - This means if the patient does not have a DVT, there is an 86% chance of the test being negative. The test will have a 14% false positive result.
- Likelihood ratio for ruling in a diagnosis.  $LR(+) = \text{sensitivity}/(1-\text{specificity})$

5.8

10%



6.9

41%

## ANSWER

- Sensitivity =  $a/(a+c) = 75/77 = 0.97 = 97\%$ 
  - This means that if the patient has a DVT, there is a 97% chance of the test being positive. The test will only have a 3% false negative rate.
- Specificity =  $d/(b+d) = 798/923 = 0.86 = 86\%$ 
  - This means if the patient does not have a DVT, there is an 86% chance of the test being negative. The test will have a 14% false positive result.
- Likelihood ratio for ruling in a diagnosis,  $LR(+) = \text{sensitivity}/(1-\text{specificity}) = 0.97/(1 - 0.86) = 6.9$ 
  - This means that a positive test result is 6.9 times more likely to be seen in a patient who has a DVT than in a patient who does not have a DVT.





## QUESTION 30

Regarding probability distribution, which of the following statements is CORRECT:

Distribution of data is always unimodal.

In skewed data the mean is usually equal to the mode.

In a negative skew, the mass of distribution is concentrated on the right.

The normal distribution is a sigmoid curve.

In a normal distribution the mean  $>$  median.

Regarding probability distribution, which of the following statements is CORRECT:

Distribution of data is always unimodal. 10%

In skewed data the mean is usually equal to the mode. 7%

✓ In a negative skew, the mass of distribution is concentrated on the right. 57%

✗ The normal distribution is a sigmoid curve. 17%

In a normal distribution the mean > median. 9%

## ANSWER

Distribution of data is usually unimodal (one peak) but may be bimodal (two peaks) or uniform (no peaks, each value equally likely). The normal distribution is a symmetrical bell-shaped curve. The mean, median, and mode of a normal distribution are equal. In a positive skew, the

Distribution of data is always unimodal. 10%

In skewed data the mean is usually equal to the mode. 7%

✓ In a negative skew, the mass of distribution is concentrated on the right. 57%

✗ The normal distribution is a sigmoid curve. 17%

In a normal distribution the mean > median. 9%

## ANSWER

Distribution of data is usually unimodal (one peak) but may be bimodal (two peaks) or uniform (no peaks, each value equally likely). The normal distribution is a symmetrical bell-shaped curve. The mean, median, and mode of a normal distribution are equal. In a positive skew, the right tail is longer and the mass of distribution is concentrated on the left;  $\text{mean} > \text{median} > \text{mode}$ . In a negative skew, the left tail is longer and the mass of distribution is concentrated on the right;  $\text{mean} < \text{median} < \text{mode}$ .

A randomised controlled trial (RCT) is performed where 1000 men are treated with a lipid-lowering drug, Superstatin, and 1000 given a placebo. Two years later the number of patients who have had a myocardial infarction (MI) is recorded and shown below. What is the relative risk reduction of treatment in preventing an MI:

|             | Myocardial infarction<br>Yes |
|-------------|------------------------------|
| Superstatin | 20                           |
| Placebo     | 40                           |
| Total       | 60                           |

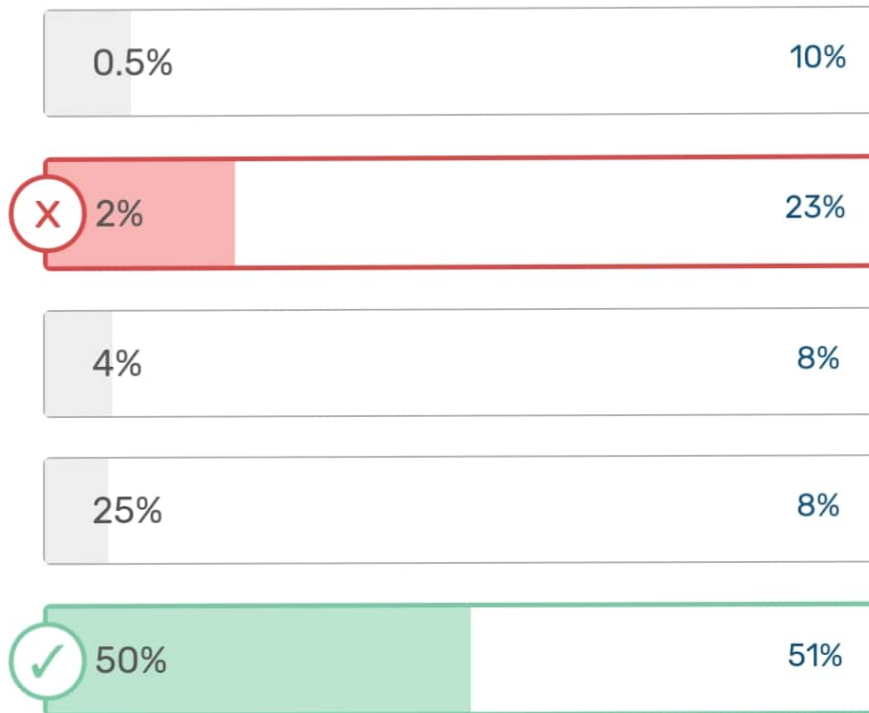
0.5%

2%

4%

25%

50%



## ANSWER

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the absolute risk of having an MI by 2%.
- The RR =  $ART/ARC = (20/1000)/(40/1000) = 0.02/0.04 = 0.5$  (50%)
  - therefore the risk of having an MI



## ANSWER

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the absolute risk of having an MI by 2%.
- The RR =  $ART/ARC = (20/1000)/(40/1000) = 0.02/0.04 = 0.5$  (50%)
  - therefore the risk of having an MI while taking Superstatin is 50% that of having an MI while not taking Superstatin.
- The RRR =  $1 - RR = 1 - 0.5 = 0.5$  (50%)
  - therefore Superstatin reduces the relative risk of having an MI by 50%.

## ANSWER

---

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the absolute risk of having an MI by 2%.
- The RR =  $ART/ARC = (20/1000)/(40/1000) = 0.02/0.04 = 0.5$  (50%)
  - therefore the risk of having an MI while taking Superstatin is 50% that of having an MI while not taking Superstatin.
- The RRR =  $1 - RR = 1 - 0.5 = 0.5$  (50%)
  - therefore Superstatin reduces the relative risk of having an MI by 50%.



Save

End Session



## QUESTION 32

Which of the following is NOT an advantage of a cohort study used to investigate the relationship between exposure to a risk factor and a future outcome:

Can be used to assess time sequence of events

Can give a direct estimate of disease incidence rates

Can measure absolute and relative risk directly

Particularly suitable for rare diseases

Can study a wide range of disease outcomes

incidence rates



Can measure absolute and relative risk directly

14%



Particularly suitable for rare diseases

64%

Can study a wide range of disease outcomes

9%

## ANSWER

### Advantages:

- ideal for studying associations between an exposure and an outcome when the exposure is uncommon
- the time sequence of events can be assessed
- they can provide information on a wide range of disease outcomes
- the absolute and relative risk of disease can be measured directly
- they can give a direct estimation of disease incidence rates

### Disadvantages:

- costly and can take long periods of time if the outcome is delayed
- subject to subject-selection and loss to follow-up bias
- large sample size required for rare



Particularly suitable for rare diseases 64%

Can study a wide range of disease outcomes 9%

## ANSWER

### Advantages:

- ideal for studying associations between an exposure and an outcome when the exposure is uncommon
- the time sequence of events can be assessed
- they can provide information on a wide range of disease outcomes
- the absolute and relative risk of disease can be measured directly
- they can give a direct estimation of disease incidence rates

### Disadvantages:

- costly and can take long periods of time if the outcome is delayed
- subject to subject-selection and loss to follow-up bias
- large sample size required for rare outcome of interest so it is not useful for rare diseases
- prone to confounding



### QUESTION 33

What is the mean of the following data set: 13, 18, 13, 14, 13, 16, 14, 21, 13



See Answer

Save

End Session

 QUESTION 33

What is the mean of the following data set: 13, 18, 13, 14, 13, 16, 14, 21, 13

- ☒ 13 14%
- ☐ 14 11%
- ☒ 15 67%
- ☐ 16 5%
- ☐ 17 3%

**ANSWER**

The mean =  $(13 + 18 + 13 + 14 + 13 + 16 + 14 + 21 + 13)/9 = 15$ .





What is the median of the following data set; 3, 13, 7, 5, 21, 24, 23, 40, 25, 21, 12, 56, 26, 29, 39



See Answer

Save

End Session



What is the median of the following data set; 3, 13, 7, 5, 21, 24, 23, 40, 25, 21, 12, 56, 26, 29, 39

20

1%

21

12%

22

17%



23

62%

25

8%

## ANSWER

When we put those numbers in order we have: 3, 5, 7, 12, 13, 21, 21, 23, 24, 25, 26, 29, 39, 40, 56.

There are fifteen numbers, so  $n = 15$ . Our middle is the eighth number  $(n+1)/2$ , therefore the median is 23.





## QUESTION 35

Regarding linear relationships between two variables, what does the correlation coefficient  $r = 1$  indicate:

The two variables are inversely proportional

The correlation is not statistically significant

There is no correlation between two variables

There is perfect correlation between two variables

One variable causes change in the other variable




See Answer



Regarding linear relationships between two variables, what does the correlation coefficient  $r = 1$  indicate:

The two variables are inversely proportional 5%

The correlation is not statistically significant 5%

 There is no correlation between two variables 11%

 There is perfect correlation between two variables 72%

One variable causes change in the other variable 7%

## ANSWER

If there is a perfect relationship between the two variables, then  $r = 1$  (+ or -). The closer that  $r$  is to 1, the greater the strength of correlation (and the closer the points are to a straight line).

correlation coefficient  $r = 1$  indicate:

The two variables are inversely proportional

5%

The correlation is not statistically significant

5%



There is no correlation between two variables

11%



There is perfect correlation between two variables

72%

One variable causes change in the other variable

7%

## ANSWER

If there is a perfect relationship between the two variables, then  $r = 1$  (+ or -). The closer that  $r$  is to 1, the greater the strength of correlation (and the closer the points are to a straight line).

Correlation does not give information about cause and effect.





A new blood test is being developed to diagnose DVT. 1000 people presenting to ED with suspected DVT undergo the new blood test and the gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the positive predictive value of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

86%

95%

98%

undergo the new blood test and the gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the positive predictive value of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

86%

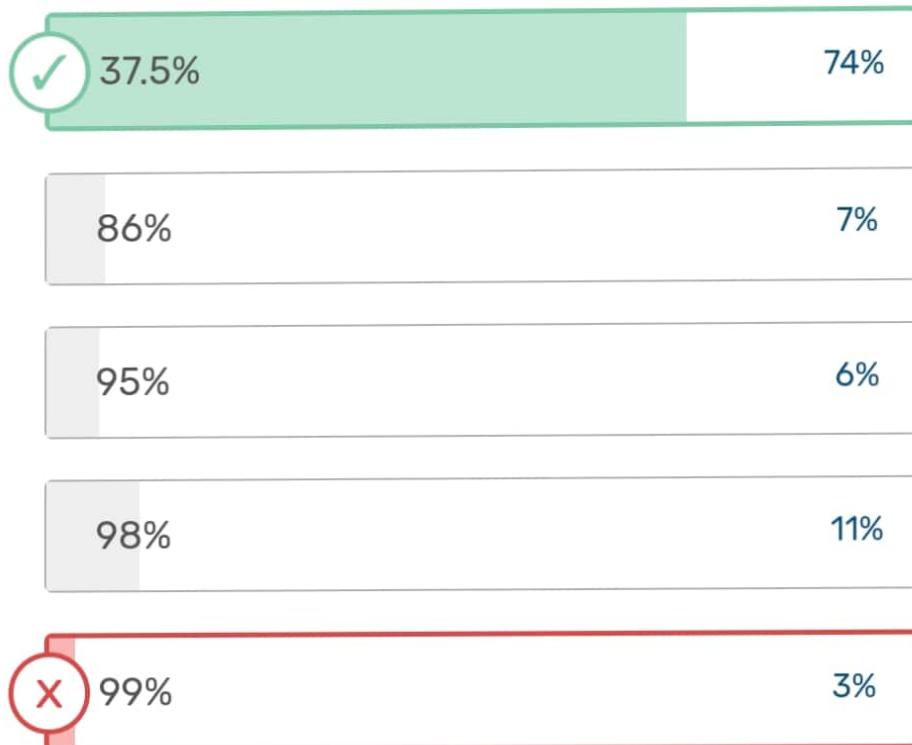
95%

98%

99%

with DVT, 120 test positive with the new diagnostic test. What is the positive predictive value of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |



## ANSWER

- Positive predictive value (PPV) =  $a/(a+b) = 75/200 = 0.375 = 37.5\%$ 
  - This means there is a 37.5% chance, if the test is positive, that the patient actually has a DVT.

 QUESTION 37

Kaplan Meier plots:

Describe the distribution and skew of data

Demonstrate the existence of publication bias in meta-analysis

Plot the probability of survival against time from intervention

Demonstrate strength of evidence of constituent trials in a meta-analysis

Demonstrate the degree of correlation between two variables





## QUESTION 37

Kaplan Meier plots:

Describe the distribution and skew of data 9%



Demonstrate the existence of publication bias in meta-analysis 11%



Plot the probability of survival against time from intervention 50%

Demonstrate strength of evidence of constituent trials in a meta-analysis 18%

Demonstrate the degree of correlation between two variables 13%

## ANSWER

The Kaplan-Meier survival curve plots the probability of survival against time from a certain start point e.g. intervention, entry into study.



## QUESTION 38

Regarding the mean, which of the following statements is INCORRECT:

The mean is calculated by summing all of the values of a data set and dividing this by the number of observations in the data set.

The mean is typically used to compare outcomes in normally distributed data.

The mean is not affected by outliers.

The mean is calculated using all of the data values.

In a normal distribution the mean is equal to the median.



Regarding the mean, which of the following statements is INCORRECT:

The mean is calculated by summing all of the values of a data set and dividing this by the number of observations in the data set. 8%

The mean is typically used to compare outcomes in normally distributed data. 9%



The mean is not affected by outliers. 72%

The mean is calculated using all of the data values. 4%

In a normal distribution the mean is equal to the median. 8%

## ANSWER

The mean is the arithmetic average, calculated by summing all of the values of a data set and dividing this sum by the number of observations in the data set.

outcomes in normally distributed data.



The mean is not affected by outliers.

72%

The mean is calculated using all of the data values.

4%

In a normal distribution the mean is equal to the median.

8%

## ANSWER

The mean is the arithmetic average, calculated by summing all of the values of a data set and dividing this sum by the number of observations in the data set.

The mean is used in normally distributed data as it uses all of the data set and roughly reflects the sampling distribution. However the mean is distorted by outliers and skewed data, where the median should be used instead.



Save

End Session



 QUESTION 39

Regarding forest plots in meta-analysis, the size of each square represents:

The 95% confidence intervals for the effect size of a study.

The effect size of a study.

The weight given to a study.

The 95% confidence interval for the summary effect size.

The summary effect size.





## QUESTION 39

Regarding forest plots in meta-analysis, the size of each square represents:

The 95% confidence intervals for the effect size of a study. 14%

The effect size of a study. 23%

✓ The weight given to a study. 41%

The 95% confidence interval for the summary effect size. 13%

The summary effect size. 9%

## ANSWER

The right-hand column is a plot of the measure of effect size, for example an odds ratio (OR), for each of these studies, represented by a square:

- The midpoint of the square represents the effect size.
- The width of the horizontal lines extending from the square represents the 95% confidence interval for this effect size.

The 95% confidence intervals for the effect size of a study. 14%

The effect size of a study. 23%

✓ The weight given to a study. 41%

The 95% confidence interval for the summary effect size. 13%

The summary effect size. 9%

## ANSWER

The right-hand column is a plot of the measure of effect size, for example an odds ratio (OR), for each of these studies, represented by a square:

- The midpoint of the square represents the effect size.
- The width of the horizontal lines extending from the square represents the 95% confidence interval for this effect size.
- The area of each square is proportional to the weight given to that study in the meta-analysis (studies with larger sample sizes, and more effect sizes are given more weight).



## QUESTION 40

Which of the following best describes an intention to treat analysis:

All patients are included in the analysis according to the group into which they were randomised only if they remain in the study.

All patients are included in the analysis according to the group into which they were randomised even if they are withdrawn from the study.

Only patients who comply with treatment are included in the analysis.

Only patients who were not lost to follow-up are included in the analysis.

Only patients who have an effect

Which of the following best describes an intention to treat analysis:

All patients are included in the analysis according to the group into which they were randomised only if they remain in the study.

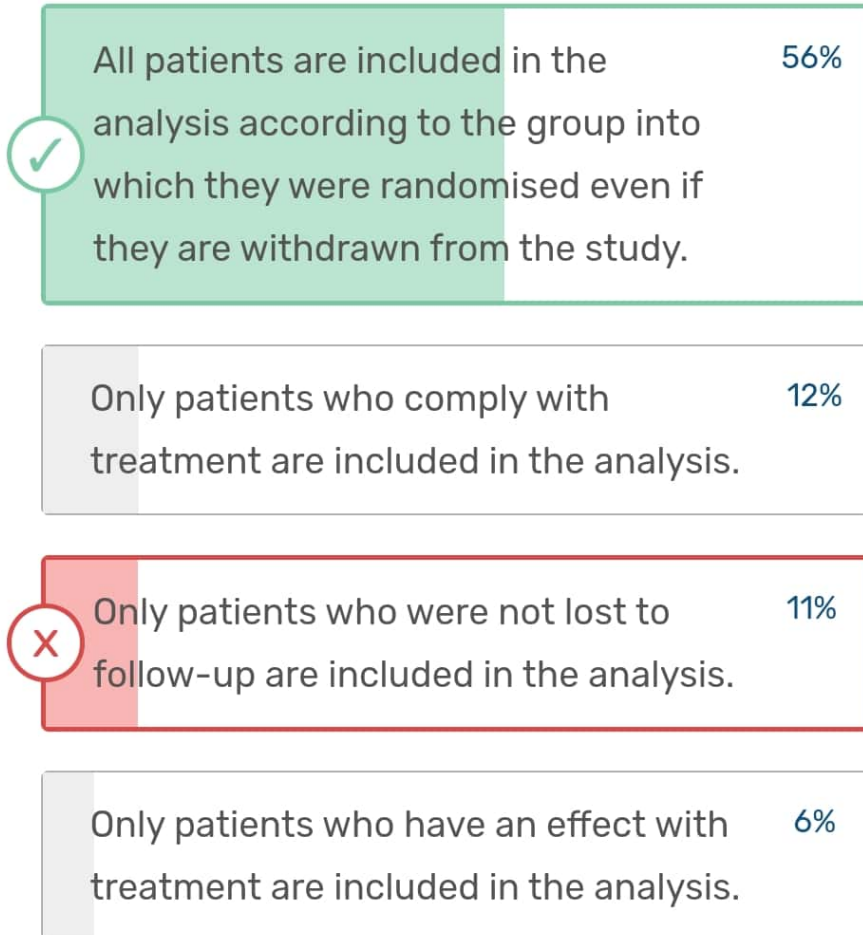
All patients are included in the analysis according to the group into which they were randomised even if they are withdrawn from the study.

Only patients who comply with treatment are included in the analysis.

Only patients who were not lost to follow-up are included in the analysis.

Only patients who have an effect with treatment are included in the analysis.





## ANSWER

An intention to treat (ITT) analysis is one in which all patients are included in the analysis, classified according to the group into which they were randomised, even if they were withdrawn from the study and did not actually receive the treatment, did not comply with treatment or drop-out. Intention to treat analysis is a more reliable estimate of true treatment effectiveness by replicating what happens in the 'real world' (e.g. noncompliance and protocol violations commonly affect therapies).

 QUESTION 41

What is the mean of the following data set; 3, 7, 5, 13, 20, 23, 39, 23, 40, 23, 14, 12, 56, 23, 29:

[See Answer](#)



## QUESTION 41

What is the mean of the following data set; 3, 7, 5, 13, 20, 23, 39, 23, 40, 23, 14, 12, 56, 23, 29:

18

3%



22

71%

25

14%

28

9%

34

3%

## ANSWER

The sum of these numbers is 330, there are fifteen numbers, therefore the mean =  $330/15 = 22$ .



## QUESTION 42

What is the range of the following data set; 3, 7, 5, 13, 20, 23, 39, 23, 40, 23, 14, 12, 56, 23, 29:

[See Answer](#)



## QUESTION 42

What is the range of the following data set; 3, 7, 5, 13, 20, 23, 39, 23, 40, 23, 14, 12, 56, 23, 29:

25

5%

40

5%

 48

6%

 53

73%

56

11%

## ANSWER

The range is the difference between the largest and smallest observations in the data set and can be given as two numbers or a single value. In this case the range is 3 - 56 or 53.



## QUESTION 43

Which of the following is an example of ordinal data:

Disease staging system

Number of children

Height

Marital status

Gender



See Answer

Save

End Session



## QUESTION 43

Which of the following is an example of ordinal data:

- |  |                        |     |
|--|------------------------|-----|
|  | Disease staging system | 67% |
|  | Number of children     | 10% |
|  | Height                 | 11% |
|  | Marital status         | 7%  |
|  | Gender                 | 6%  |

## ANSWER

Ordinal data is categorical data that is ordered in some way e.g. disease staging system, pain scoring system.



 QUESTION 44

The power of a study is:

The probability of correctly accepting the null hypothesis when it is true.

The probability of correctly rejecting the null hypothesis when it is false.

The probability of getting clinically significant results.

The probability of wrongly rejecting the null hypothesis when it is true.

The probability of wrongly accepting the null hypothesis when it is false.





## QUESTION 44

The power of a study is:

The probability of correctly accepting the null hypothesis when it is true. 15%



The probability of correctly rejecting the null hypothesis when it is false. 45%

The probability of getting clinically significant results. 29%



The probability of wrongly rejecting the null hypothesis when it is true. 6%

The probability of wrongly accepting the null hypothesis when it is false. 6%

## ANSWER

The power of a study is the probability of correctly rejecting the null hypothesis when it is false. It is the probability that the study would detect a statistically significant difference if a true difference exists. As the power increases, the chance of making a type II error decreases.

The probability of correctly accepting the null hypothesis when it is true. 15%



The probability of correctly rejecting the null hypothesis when it is false. 45%

The probability of getting clinically significant results. 29%



The probability of wrongly rejecting the null hypothesis when it is true. 6%

The probability of wrongly accepting the null hypothesis when it is false. 6%

## ANSWER

The power of a study is the probability of correctly rejecting the null hypothesis when it is false. It is the probability that the study would detect a statistically significant difference if a true difference exists. As the power increases, the chance of making a type II error decreases.



Save

End Session



A new test is being developed to diagnose chlamydia. 1000 people aged 15 – 35 years attending a GUM clinic undergo the new test and the current gold standard nucleic acid amplification test (NAAT) to confirm the diagnosis. Of the 1000 people, 250 are diagnosed with chlamydia. Of the patients diagnosed with chlamydia, 240 test positive with the new diagnostic test and of the patients not diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the positive predictive value of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |

amplification test (NAAT) to confirm the diagnosis. Of the 1000 people, 250 are diagnosed with chlamydia. Of the patients diagnosed with chlamydia, 240 test positive with the new diagnostic test and of the patients not diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the positive predictive value of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |

62%

75%

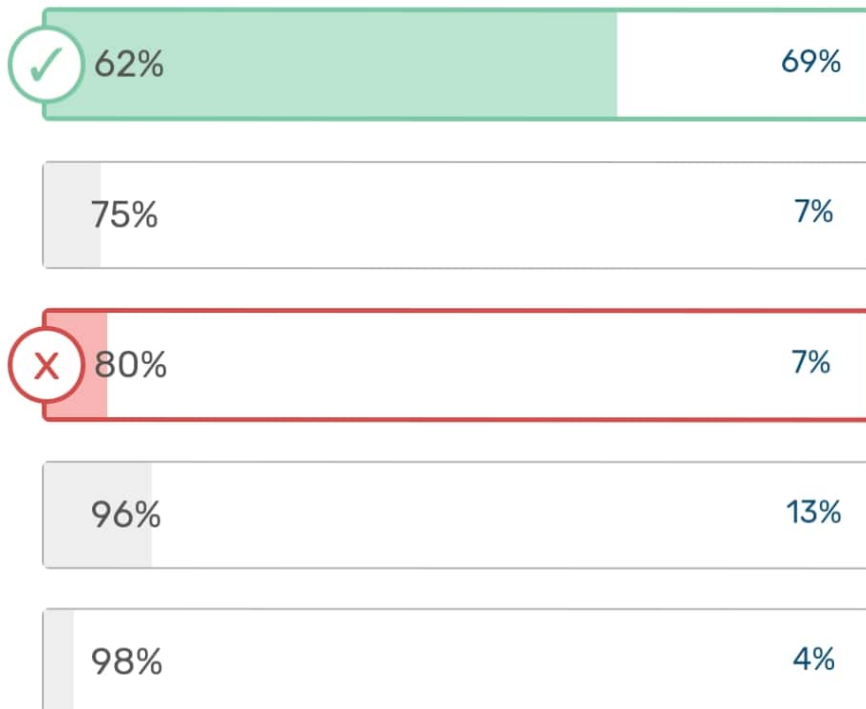
80%

96%

98%

test. What is the positive predictive value of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |



## ANSWER

- Positive predictive value (PPV) =  $a/(a+b)$  =  $240/390 = 0.62 = 62\%$ 
  - This means there is a 62% chance, if the test is positive, that the patient actually has chlamydia.



## QUESTION 46

Which of the following statistical tests is most appropriate for comparing means between two independent parametric variables:

Mann Whitney U-test

Kruskal-Wallis test

Wilcoxon matched pairs test

Unpaired student t-test

Chi-squared test



See Answer

Which of the following statistical tests is most appropriate for comparing means between two independent parametric variables:

Mann Whitney U-test

8%

Kruskal-Wallis test

4%



Wilcoxon matched pairs test

6%



Unpaired student t-test

59%

Chi-squared test

24%

## ANSWER

The unpaired student t-test would be most appropriate to compare the means of two independent groups of values.



Save

End Session



If the null hypothesis is wrongly rejected when it is actually true, this is an example of:

A test with high power

A test with a high negative predictive value

A test with a type I error

A test with nonparametric data

A test with a type II error



See Answer

Save

End Session

SCORE

20%



If the null hypothesis is wrongly rejected when it is actually true, this is an example of:

A test with high power

2%

A test with a high negative predictive value

4%



A test with a type I error

69%



A test with nonparametric data

2%

A test with a type II error

24%

## ANSWER

A type I error occurs when the null hypothesis is wrongly rejected when it is actually true and we conclude that there is a difference of effect when in reality there is none (a false positive result).



Save

End Session



A new blood test is being developed to diagnose DVT. 1000 people presenting to ED with suspected DVT undergo the new blood test and the gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the sensitivity of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

86%

95%

97%

gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the sensitivity of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a = 75  |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

86%

95%

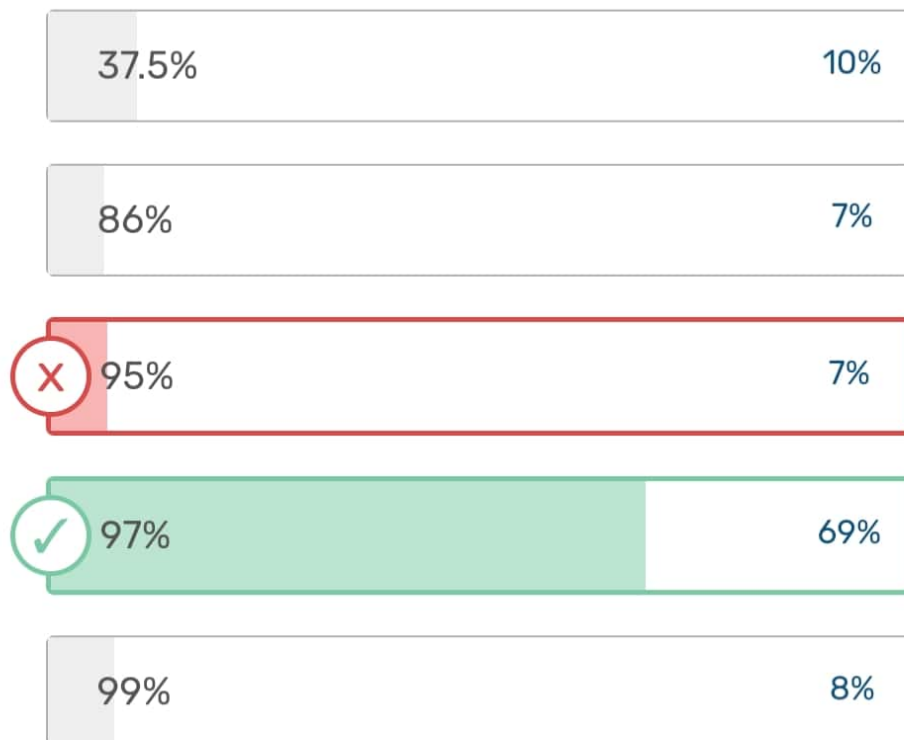
97%

99%



with DVT, 125 test positive with the new diagnostic test. What is the sensitivity of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |



## ANSWER

- Sensitivity =  $a/(a+c) = 75/77 = 0.97 = 97\%$ 
  - This means that if the patient has a DVT, there is a 97% chance of the test being positive. The test will only have a 3% false negative rate.

 QUESTION 49

What is the interquartile range for the following data set; 11, 4, 6, 8, 3, 10, 8, 10, 4, 12, 31

[See Answer](#)



## QUESTION 49

What is the interquartile range for the following data set; 11, 4, 6, 8, 3, 10, 8, 10, 4, 12, 31

5

9%



7

31%



8

26%

10

17%

28

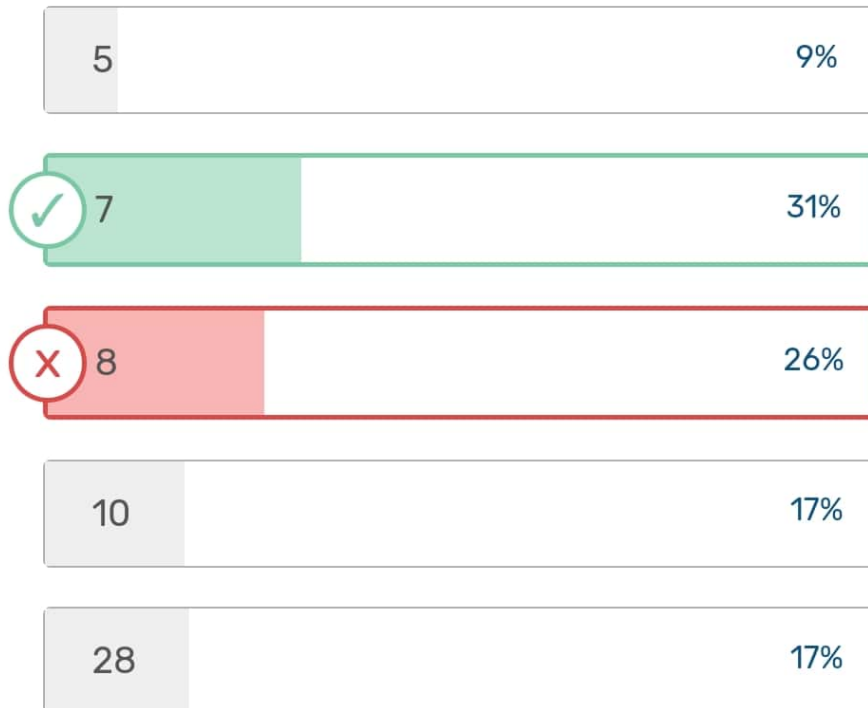
17%

## ANSWER

Ordering the data, we get 3, 4, 4, 6, 8, 8, 10, 10, 11, 12 and 31, and sample size  $(n) = 11$ .

Median =  $[(n+1)/2]$ th value =  $(11+1)/2 = 6$ th value  
= 8

What is the interquartile range for the following data set; 11, 4, 6, 8, 3, 10, 8, 10, 4, 12, 31



## ANSWER

Ordering the data, we get 3, 4, 4, 6, 8, 8, 10, 10, 11, 12 and 31, and sample size  $(n) = 11$ .

Median =  $[(n+1)/2]$ th value =  $(11+1)/2 = 6$ th value  
= 8

The lower (first) quartile =  $[(n+1)/4]$ th value =  $(11 + 1)/4 = 3$ rd value = 4

The upper (third) quartile =  $3[(n+1)/4]$ th value =  $3[(11 + 1)/4] = 9$ th value = 11.

The interquartile range is the difference between



## ANSWER

Ordering the data, we get 3, 4, 4, 6, 8, 8, 10, 10, 11, 12 and 31, and sample size  $(n) = 11$ .

Median =  $[(n+1)/2]$ th value =  $(11+1)/2 = 6$ th value  
= 8

The lower (first) quartile =  $[(n+1)/4]$ th value =  $(11 + 1)/4 = 3$ rd value = 4

The upper (third) quartile =  $3[(n+1)/4]$ th value =  $3[(11 + 1)/4] = 9$ th value = 11.

The interquartile range is the difference between the upper quartile and lower quartile =  $11 - 4 = 7$ .



Save

End Session





## QUESTION 50

You are presenting at a journal club and critically appraising a research paper based on a randomised control trial. How can researchers reduce the chance of making a type II error:

Increase accuracy of measuring tools

Decrease the significance level

Decrease confidence intervals

Increase the sample size

Ensure double blinding is used





## QUESTION 50

You are presenting at a journal club and critically appraising a research paper based on a randomised control trial. How can researchers reduce the chance of making a type II error:

Increase accuracy of measuring tools 12%

Decrease the significance level 0%

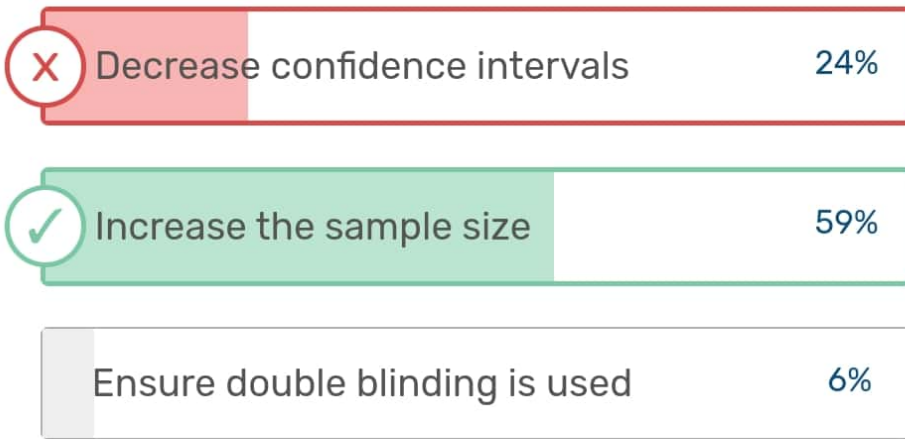
☒ Decrease confidence intervals 24%

☒ Increase the sample size 59%

Ensure double blinding is used 6%

## ANSWER

A type II error occurs when the null hypothesis is wrongly accepted when it is actually false and we conclude that there is no evidence of a



## ANSWER

A type II error occurs when the null hypothesis is wrongly accepted when it is actually false and we conclude that there is no evidence of a difference in effect when one really exists (a false negative result).

The probability of making a type II error is denoted by beta ( $\beta$ ).

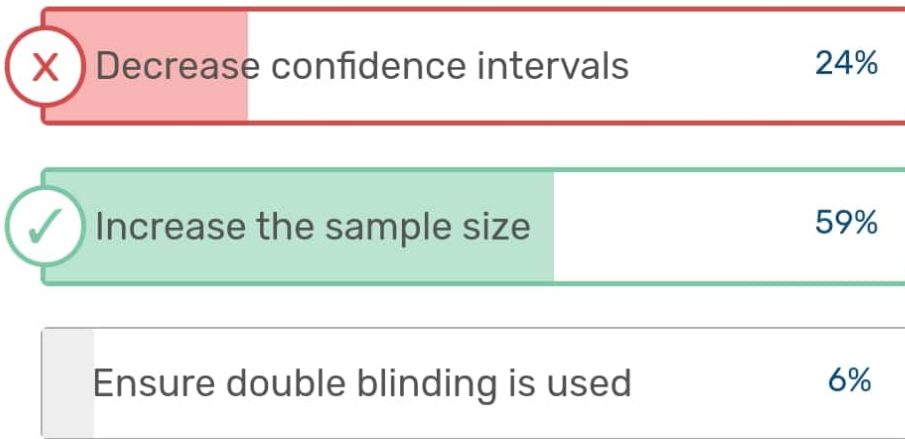
Type II errors can be avoided at the design stage of the study by power calculations that give an indication of the minimum sample size required so that one can be reasonable likely to detect an effect of a given size.



Save

End Session





## ANSWER

A type II error occurs when the null hypothesis is wrongly accepted when it is actually false and we conclude that there is no evidence of a difference in effect when one really exists (a false negative result).

The probability of making a type II error is denoted by beta ( $\beta$ ).

Type II errors can be avoided at the design stage of the study by power calculations that give an indication of the minimum sample size required so that one can be reasonable likely to detect an effect of a given size.



Save

End Session





## QUESTION 51

What is the mode of the following data set; 3, 7, 5, 13, 20, 23, 39, 23, 40, 23, 14, 12, 56, 23, 29:

[See Answer](#)



## QUESTION 51

What is the mode of the following data set; 3, 7, 5, 13, 20, 23, 39, 23, 40, 23, 14, 12, 56, 23, 29:

20

1%



21

2%

22

5%



23

90%

25

2%

## ANSWER

In order these numbers are: 3, 5, 7, 12, 13, 14, 20, **23, 23, 23, 23**, 29, 39, 40, 56. This makes it easy to see which numbers appear most often. In this case the mode is 23.



An FY2 in ED is collecting retrospective data for an audit on headache. He records the number of patients with headaches seen per day and for each patient who presented to ED with a headache he collects data on the length of time waiting to see the triage nurse, the pain score at triage (Faces Pain Rating Scale) and pain relief given in the ED (e.g. simple non-opioid analgesia, opioid analgesia and nil).

The pain-scoring system given is an example of which of the following types of data:

Qualitative

Nominal

Ordinal

Discrete

Continuous

patient who presented to ED with a headache he collects data on the length of time waiting to see the triage nurse, the pain score at triage (Faces Pain Rating Scale) and pain relief given in the ED (e.g. simple non-opioid analgesia, opioid analgesia and nil). The pain-scoring system given is an example of which of the following types of data:

|   |     |
|---|-----|
| Qualitative                                 | 15% |
| Nominal                                     | 14% |
| <input checked="" type="checkbox"/> Ordinal | 53% |
| <input type="checkbox"/> Discrete           | 13% |
| Continuous                                  | 5%  |

## ANSWER

A pain scoring system is an example of ordinal categorical data.

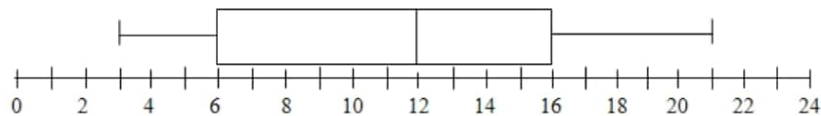


Save

End Session

 QUESTION 53

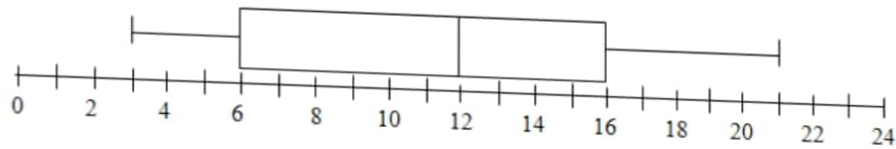
Regarding the following box and whisker plot, what percentage of data values lie between 6 and 16.





## QUESTION 53

Regarding the following box and whisker plot, what percentage of data values lie between 6 and 16.



25%

4%



50%

67%



75%

23%

100%

5%

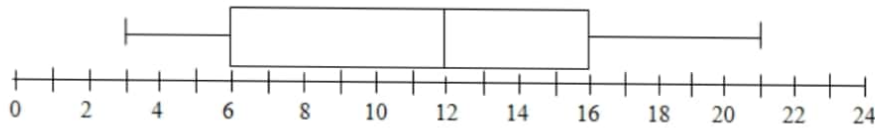
0%

1%

## ANSWER

The rectangle drawn between 6 and 16 contains 50% of the data values. The interquartile range (IQ) is a measure of spread given by the difference between the first quartile (the value below which 25% of the observations lie) and the third quartile (the value below which

Whisker plot, what percentage of data values lie between 6 and 16.



25%

4%



50%

67%



75%

23%

100%

5%

0%

1%

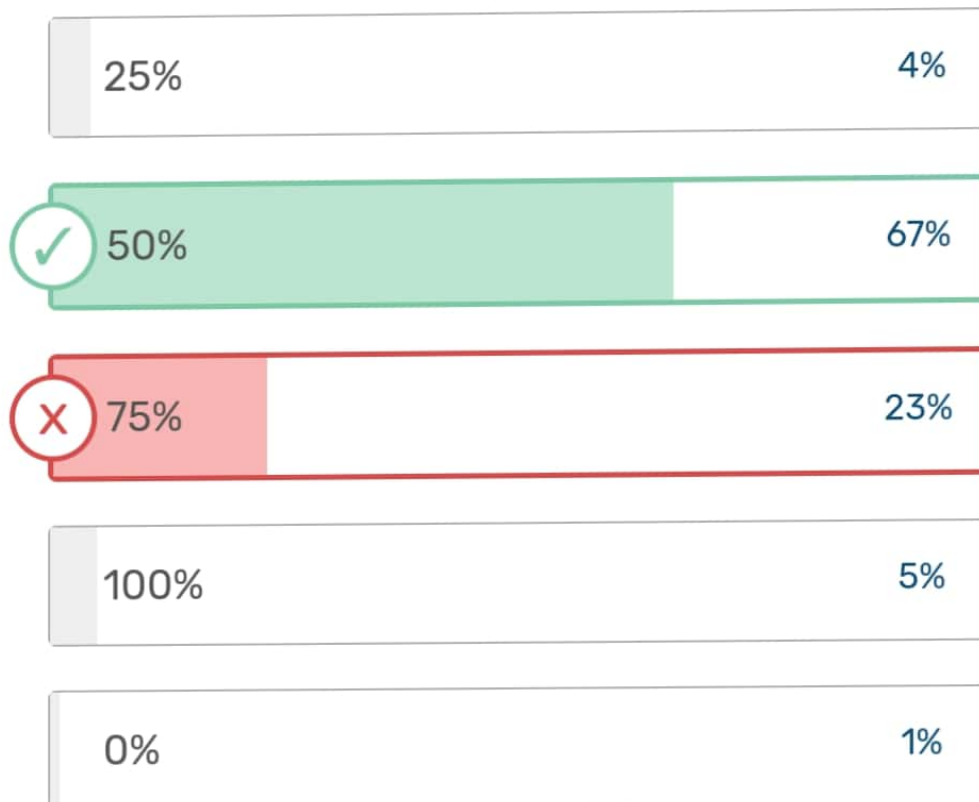
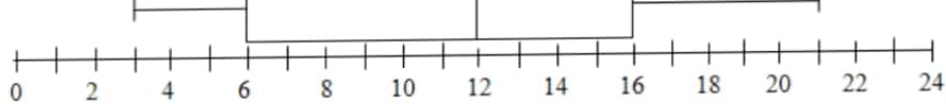
## ANSWER

The rectangle drawn between 6 and 16 contains 50% of the data values. The interquartile range (IQ) is a measure of spread given by the difference between the first quartile (the value below which 25% of the observations lie) and the third quartile (the value below which 75% of the observations lie).



Save

End Session



## ANSWER

The rectangle drawn between 6 and 16 contains 50% of the data values. The interquartile range (IQ) is a measure of spread given by the difference between the first quartile (the value below which 25% of the observations lie) and the third quartile (the value below which 75% of the observations lie).



Save

End Session



QUESTION 54

Which of the following is most true of normally distributed data:

Mean > Median

Median > Mean

Mean = Median = Mode

Mean > Mode

Median < Mode



See Answer

Save

End Session



## QUESTION 54

Which of the following is most true of normally distributed data:

Mean > Median

6%



Median > Mean

4%



Mean = Median = Mode

86%

Mean > Mode

3%

Median < Mode

2%

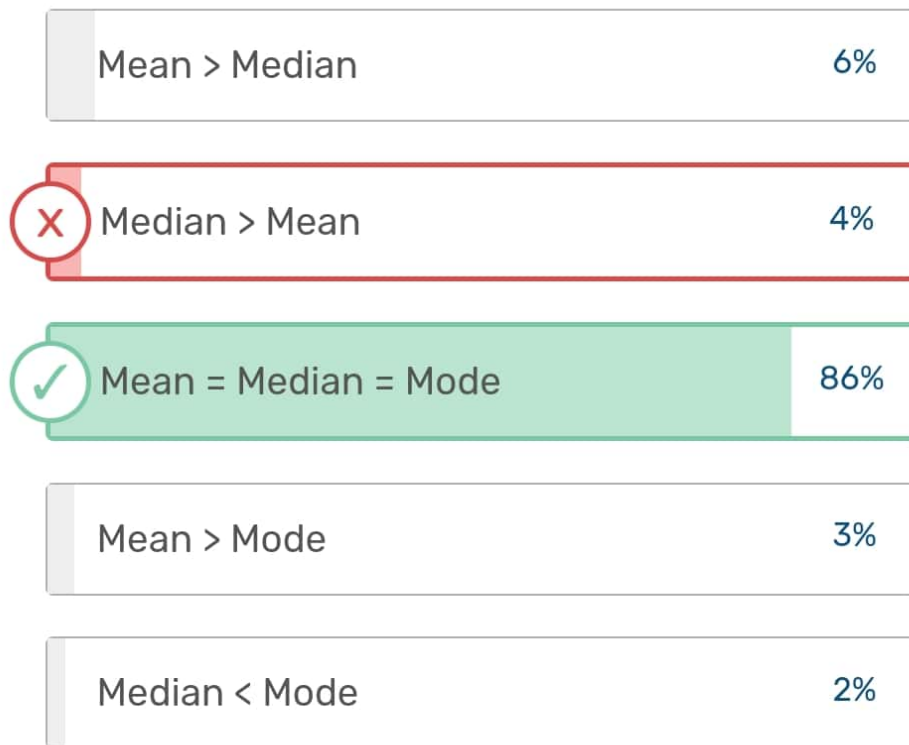
## ANSWER

The mean, median, and mode of a normal distribution are equal.

In a positive skew: mean > median > mode.

In a negative skew: mean < median < mode

normally distributed data:



## ANSWER

The mean, median, and mode of a normal distribution are equal.

In a positive skew:  $\text{mean} > \text{median} > \text{mode}$ .

In a negative skew:  $\text{mean} < \text{median} < \text{mode}$



Save

End Session



SCORE

17%



## QUESTION 55

Regarding forest plots in meta-analysis, the lateral points of the diamond represents:

The 95% confidence interval for the effect size of a study.

The effect size of a study.

The weight given to a study.

The 95% confidence interval for the summary effect size.

The summary effect size.





## QUESTION 55

Regarding forest plots in meta-analysis, the lateral points of the diamond represents:

The 95% confidence interval for the effect size of a study. 20%

The effect size of a study. 15%

✗ The weight given to a study. 8%

✓ The 95% confidence interval for the summary effect size. 45%

The summary effect size. 12%

## ANSWER

The overall meta-analysed summary effect size is plotted as a diamond:

- The midpoint of the diamond represents the summary effect size.
- The lateral points represent the 95% confidence interval for this estimate.
- There may be an ascending dotted line

The 95% confidence interval for the effect size of a study. 20%

The effect size of a study. 15%

✗ The weight given to a study. 8%

✓ The 95% confidence interval for the summary effect size. 45%

The summary effect size. 12%

## ANSWER

The overall meta-analysed summary effect size is plotted as a diamond:

- The midpoint of the diamond represents the summary effect size.
- The lateral points represent the 95% confidence interval for this estimate.
- There may be an ascending dotted line from the upper point of the diamond.



Save

End Session

A randomised controlled trial (RCT) is carried out to study the effects of a new anticoagulant, Anticlot, on stroke incidence in patients with atrial fibrillation (AF). 500 patients with AF are treated with Anticlot and 500 receive usual therapy (control group). One year later the number of patients who have had a stroke is recorded and shown below. What is the relative risk of having a stroke in patients in the treatment group compared to the control group:

|          | Stroke Yes |
|----------|------------|
| Anticlot | 30         |
| Control  | 50         |
| Total    | 80         |

4%

6%

10%

100%

fibrillation (AF). 500 patients with AF are treated with Anticlot and 500 receive usual therapy (control group). One year later the number of patients who have had a stroke is recorded and shown below. What is the relative risk of having a stroke in patients in the treatment group compared to the control group:

|          | Stroke Yes |
|----------|------------|
| Anticlot | 30         |
| Control  | 50         |
| Total    | 80         |

4%

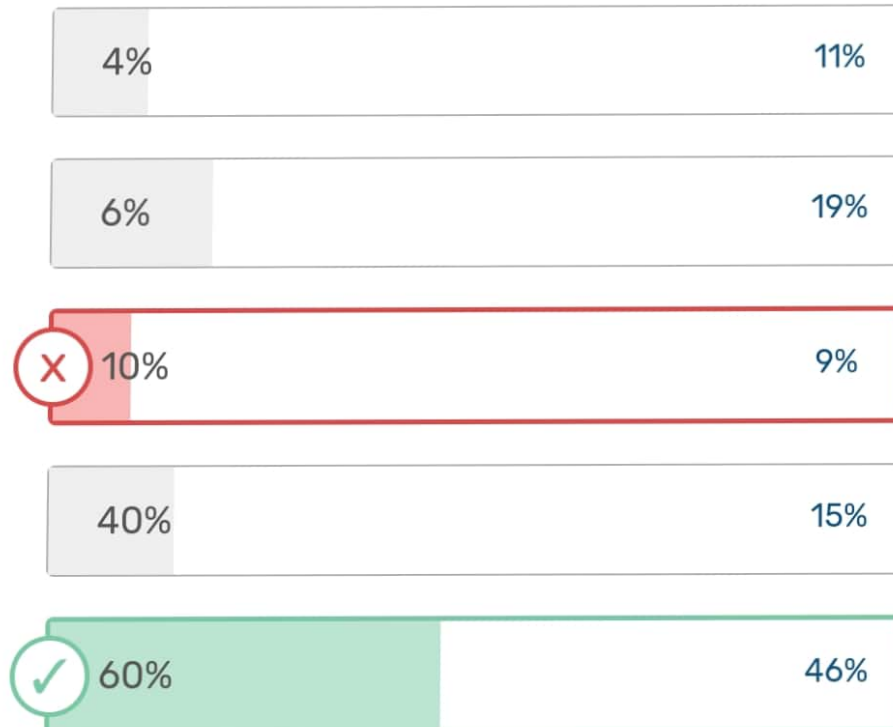
6%

10%

40%

60%





## ANSWER

- The AR of having a stroke in the treatment group (ART) =  $30/500 = 0.06$  (6%)
  - therefore the risk of having a stroke while taking Anticlot is 6%.
- The AR of having a stroke in the control group (ARC) =  $50/500 = 0.1$  (10%)
  - therefore the risk of having a stroke while not taking Anticlot is 10%.
- The ARR =  $ARC - ART = (50/500) - (30/500) = 0.10 - 0.06 = 0.04$  (4%)
  - therefore Anticlot reduces the absolute risk of having a stroke by 4%.
- The RR =  $ART/ARC = (30/500)/(50/500) = 0.06/0.1 = 0.6$  (60%)
  - therefore the risk of having a stroke

## ANSWER

---

- The AR of having a stroke in the treatment group (ART) =  $30/500 = 0.06$  (6%)
  - therefore the risk of having a stroke while taking Anticlot is 6%.
- The AR of having a stroke in the control group (ARC) =  $50/500 = 0.1$  (10%)
  - therefore the risk of having a stroke while not taking Anticlot is 10%.
- The ARR =  $ARC - ART = (50/500) - (30/500) = 0.10 - 0.06 = 0.04$  (4%)
  - therefore Anticlot reduces the absolute risk of having a stroke by 4%.
- The RR =  $ART/ARC = (30/500)/(50/500) = 0.06/0.1 = 0.6$  (60%)
  - therefore the risk of having a stroke while taking Anticlot is 60% that of having a stroke while not taking Anticlot.
- The RRR =  $1 - RR = 1 - 0.6 = 0.4$  (40%)
  - therefore Anticlot reduces the relative risk of having a stroke by 40%.





## QUESTION 57

Regarding forest plots in meta-analysis, the midpoint of the diamond represents:

The 95% confidence interval for the effect size of a study.

The effect size of a study.

The weight given to a study.

The 95% confidence interval for the summary effect size.

The summary effect size.





## QUESTION 57

Regarding forest plots in meta-analysis, the midpoint of the diamond represents:

The 95% confidence interval for the effect size of a study. 11%

The effect size of a study. 17%

☒ The weight given to a study. 13%

The 95% confidence interval for the summary effect size. 18%

☒ The summary effect size. 41%

## ANSWER

The overall meta-analysed summary effect size is plotted as a diamond:

- The midpoint of the diamond represents the summary effect size.
- The lateral points represent the 95% confidence interval for this estimate.
- There may be an ascending dotted line

The 95% confidence interval for the effect size of a study.

11%

The effect size of a study.

17%



The weight given to a study.

13%

The 95% confidence interval for the summary effect size.

18%



The summary effect size.

41%

## ANSWER

The overall meta-analysed summary effect size is plotted as a diamond:

- The midpoint of the diamond represents the summary effect size.
- The lateral points represent the 95% confidence interval for this estimate.
- There may be an ascending dotted line from the upper point of the diamond.



Save

End Session



Regarding the interpretation of risk of an outcome event when comparing a treatment and control group, how is the relative risk reduction calculated:

Absolute risk in the control group -  
Absolute risk in the treatment group

Absolute risk in the control group /  
Absolute risk in the treatment group

Absolute risk in the treatment group  
- Absolute risk in the control group

Absolute risk in the treatment group  
/ Absolute risk in the control group

1 - Relative risk



See Answer

Save

End Session



Regarding the interpretation of risk of an outcome event when comparing a treatment and control group, how is the relative risk reduction calculated:

Absolute risk in the control group - Absolute risk in the treatment group 13%

Absolute risk in the control group / Absolute risk in the treatment group 13%



Absolute risk in the treatment group - Absolute risk in the control group 9%

Absolute risk in the treatment group / Absolute risk in the control group 22%



1 - Relative risk 43%

## ANSWER

The relative risk reduction (RRR) is the proportional reduction in rates of outcome event between the control group and the treatment group.

$$RRR = (ARC - ART)/ARC = 1 - RR$$

Absolute risk in the control group - 13%  
Absolute risk in the treatment group

Absolute risk in the control group / 13%  
Absolute risk in the treatment group

✗ Absolute risk in the treatment group - 9%  
Absolute risk in the control group

Absolute risk in the treatment group / 22%  
Absolute risk in the control group

✓ 1 - Relative risk 43%

## ANSWER

The relative risk reduction (RRR) is the proportional reduction in rates of outcome event between the control group and the treatment group.

$$RRR = (ARC - ART)/ARC = 1 - RR$$

Where ARC = absolute risk of outcome event in control group and ART = absolute risk of outcome event in treatment group.





## QUESTION 59

Funnel plots are typically used to display:

The median and interquartile range

The existence of publication bias in meta-analysis

The survival of a sample cohort

The strength of evidence of constituent trials in a meta-analysis

Correlation between two variables



See Answer



## QUESTION 59

Funnel plots are typically used to display:

The median and interquartile range

9%



The existence of publication bias in meta-analysis

44%



The survival of a sample cohort

11%

The strength of evidence of constituent trials in a meta-analysis

23%

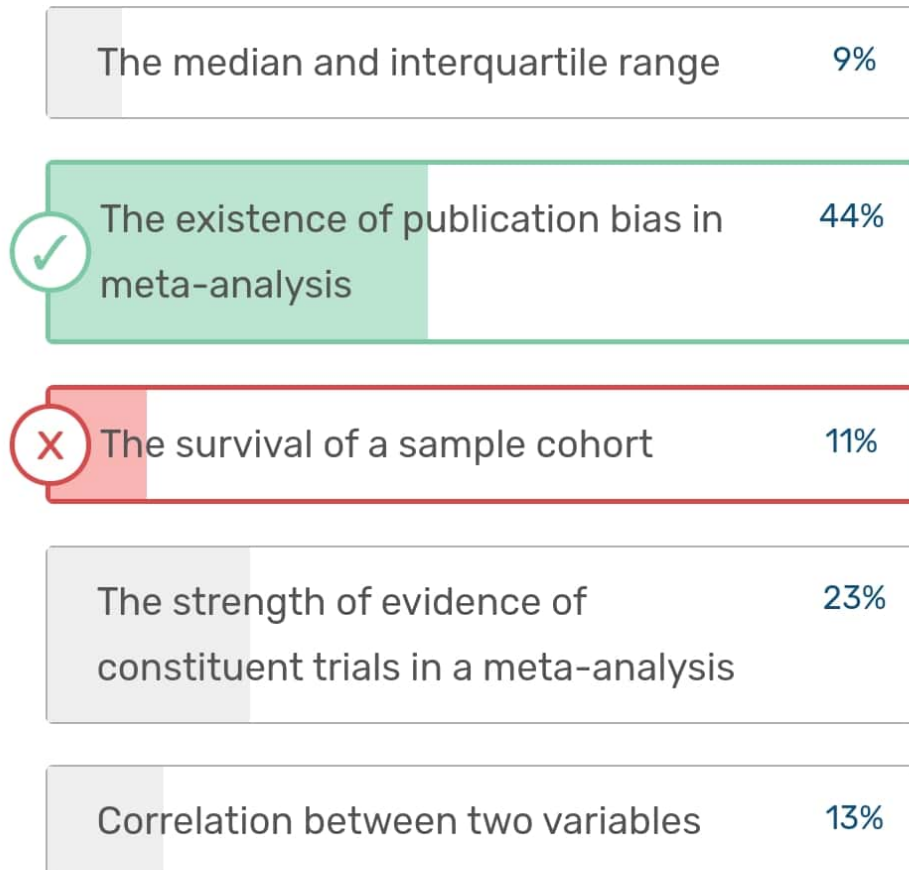
Correlation between two variables

13%

### ANSWER

Funnel plots are used to demonstrate the existence of publication bias in meta-analysis. Funnel plots are scatter plots of treatment effects estimated from individual studies on the

Funnel plots are typically used to display:



## ANSWER

Funnel plots are used to demonstrate the existence of publication bias in meta-analysis. Funnel plots are scatter plots of treatment effects estimated from individual studies on the x axis and some measure of study size on the y axis. Each point on the graph represents one of the studies. A symmetrical inverted funnel shape indicates an absence of publication bias. If there is publication bias, there will be asymmetry of the open wide end due to the absence of small negative results.



## QUESTION 60

A randomised controlled trial (RCT) is carried out to study the effects of a new anticoagulant, Anticlot, on stroke incidence in patients with atrial fibrillation (AF). 500 patients with AF are treated with Anticlot and 500 receive usual therapy (control group). One year later the number of patients who have had a stroke is recorded and shown below. What is the number needed to treat (NNT):

|          | Stroke Yes |
|----------|------------|
| Anticlot | 30         |
| Control  | 50         |
| Total    | 80         |

A randomised controlled trial (RCT) is carried out to study the effects of a new anticoagulant, Anticlot, on stroke incidence in patients with atrial fibrillation (AF). 500 patients with AF are treated with Anticlot and 500 receive usual therapy (control group). One year later the number of patients who have had a stroke is recorded and shown below. What is the number needed to treat (NNT):

|          | Stroke Yes |
|----------|------------|
| Anticlot | 30         |
| Control  | 50         |
| Total    | 80         |

20

25

40

50

60

|         |    |
|---------|----|
| Control | 50 |
| Total   | 80 |



## ANSWER

- The AR of having a stroke in the treatment group (ART) =  $30/500 = 0.06$  (6%)
  - therefore the risk of having a stroke while taking Anticlot is 6%.
- The AR of having a stroke in the control group (ARC) =  $50/500 = 0.1$  (10%)
  - therefore the risk of having a stroke while not taking Anticlot is 10%.
- The ARR =  $ARC - ART = (50/500) - (30/500) = 0.10 - 0.06 = 0.04$  (4%)
  - therefore Anticlot reduces the absolute risk of having a stroke by



## ANSWER

- The AR of having a stroke in the treatment group (ART) =  $30/500 = 0.06$  (6%)
  - therefore the risk of having a stroke while taking Anticlot is 6%.
- The AR of having a stroke in the control group (ARC) =  $50/500 = 0.1$  (10%)
  - therefore the risk of having a stroke while not taking Anticlot is 10%.
- The ARR =  $ARC - ART = (50/500) - (30/500) = 0.10 - 0.06 = 0.04$  (4%)
  - therefore Anticlot reduces the absolute risk of having a stroke by 4%.
- $NNT = 1/ARR = 1/0.04 = 25$ 
  - therefore 25 patients would need to be treated with Anticlot to prevent one extra patient from having a stroke.



Which of the following best describes the specificity of a diagnostic test:

The proportion of individuals with a positive test result who have the disease

The ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease

The proportion of patients with the disease who are correctly identified by the test as having the disease

The proportion of individuals with a negative test result who do not have the disease

The proportion of patients without the disease who are correctly identified by the test as not having the disease

disease

The ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease

4%

X

The proportion of patients with the disease who are correctly identified by the test as having the disease

12%

The proportion of individuals with a negative test result who do not have the disease

21%

✓

The proportion of patients without the disease who are correctly identified by the test as not having the disease

57%

## ANSWER

Specificity is the proportion of patients without the disease (true negatives) who are correctly identified by the test as not having the disease = true negative rate.





## QUESTION 62

A p value  $< 0.05$  obtained from a study with a significance level ( $\alpha$ ) of 0.05, means all of the following, EXCEPT:

the result is statistically significant.

the probability of obtaining the result by chance is less than 1 in 20.

the null hypothesis is rejected.

the alternative hypothesis is accepted.

the result is clinically significant.



A p value  $< 0.05$  obtained from a study with a significance level ( $\alpha$ ) of 0.05, means all of the following, EXCEPT:

the result is statistically significant.

8%

the probability of obtaining the result by chance is less than 1 in 20.

12%

the null hypothesis is rejected.

16%



the alternative hypothesis is accepted.

14%



the result is clinically significant.

50%

## ANSWER

A p value  $< 0.05$ :

- is statistically significant
- means that the probability of obtaining a given result by chance is less than 1 in 20
- means the null hypothesis is rejected
- means there is evidence of an association between a variable and an outcome

Note that this does not tell us whether the result is clinically significant.



## QUESTION 63

What is the range of the following data set: 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70

[See Answer](#)



What is the range of the following data set: 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70

50

3%

55

4%

X

60

10%

✓

65

78%

70

5%

## ANSWER

The range states the difference between the largest and smallest values in the data set and can be given as two numbers or a single value.  
The range =  $70 - 5 = 65$  (or  $5 - 75$ ).



Save

End Session



Regarding a cohort study, which of the following statements is INCORRECT:

It is prospective.

It is observational.

The usual outcome measure is the relative risk.

It is subject to loss to follow up bias.

It is useful for rare diseases.



See Answer

Save

End Session

SCORE

15%



Regarding a cohort study, which of the following statements is INCORRECT:

It is prospective.

8%

It is observational.

5%



The usual outcome measure is the relative risk.

13%

It is subject to loss to follow up bias.

11%



It is useful for rare diseases.

63%

## ANSWER

A cohort study is a longitudinal, prospective, observational study that follows a defined group (cohort) matched to unexposed controls for a set period of time and investigates the effect of exposure to a risk factor on a particular future outcome. The usual outcome measure is the relative risk (risk ratio). A large sample size is required for a rare outcome of interest so it is not useful for rare diseases.

A new blood test is being developed to diagnose DVT. 1000 people presenting to ED with suspected DVT undergo the new blood test and the gold standard doppler ultrasound to confirm the diagnosis. Of the 1000 people, 77 are confirmed to have a DVT. Of the patients diagnosed with DVT, 75 test positive with the new diagnostic test and of the patients not diagnosed with DVT, 125 test positive with the new diagnostic test. What is the specificity of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

86%

95%

98%

99%

new diagnostic test. What is the specificity of this test:

|               | DVT Yes |
|---------------|---------|
| Positive test | a= 75   |
| Negative test | c = 2   |
| Total         | 77      |

37.5%

7%



86%

64%



95%

6%

98%

8%

99%

14%

## ANSWER

- Specificity =  $d/(b+d) = 798/923 = 0.86 = 86\%$ 
  - This means if the patient does not have a DVT, there is an 86% chance of the test being negative. The test will have a 14% false positive result.

|       |    |
|-------|----|
| Total | 77 |
|-------|----|



## ANSWER

- Specificity =  $d/(b+d) = 798/923 = 0.86 = 86\%$ 
  - This means if the patient does not have a DVT, there is an 86% chance of the test being negative. The test will have a 14% false positive result.



Save

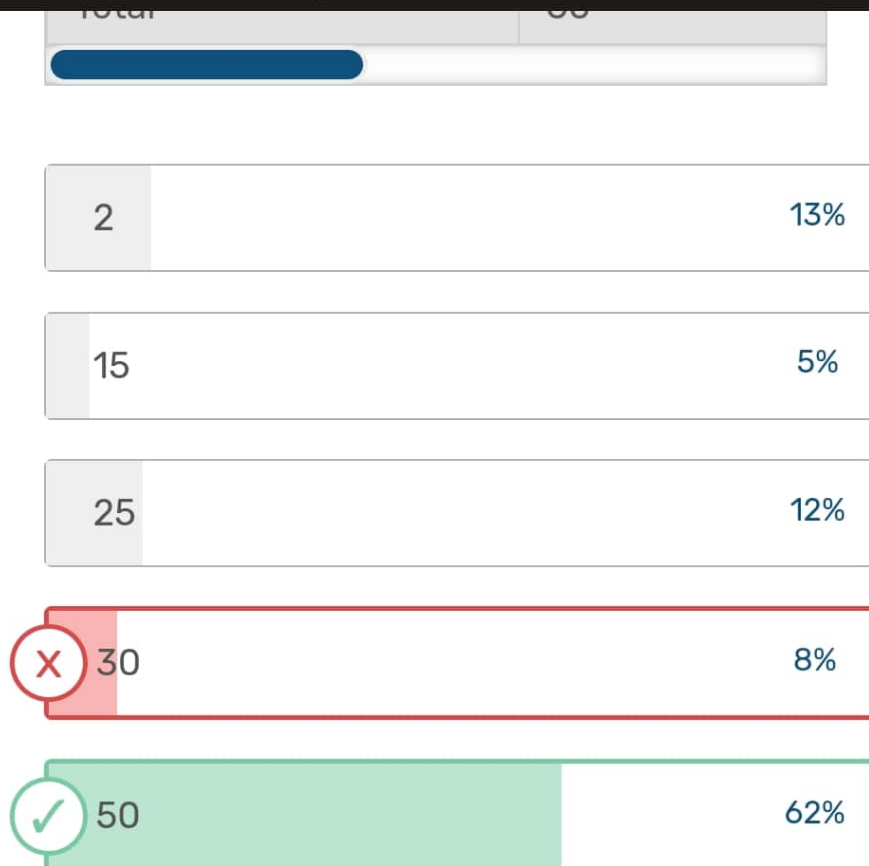
End Session





A randomised controlled trial (RCT) is performed where 1000 men are treated with a lipid-lowering drug, Superstatin, and 1000 given a placebo. Two years later the number of patients who have had a myocardial infarction (MI) is recorded and shown below. How many patients would need to be treated to prevent one extra patient from having an MI:

|             | Myocardial infarction<br>Yes |
|-------------|------------------------------|
| Superstatin | 20                           |
| Placebo     | 40                           |
| Total       | 60                           |



## ANSWER

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the absolute risk of having an MI by 2%.
- NNT =  $1/ARR = 1/0.02 = 50$  people
  - therefore 50 patients would need to



## ANSWER

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the absolute risk of having an MI by 2%.
- NNT =  $1/ARR = 1/0.02 = 50$  people
  - therefore 50 patients would need to be treated with Superstatin to prevent one extra patient from having an MI.

A new test is being developed to diagnose chlamydia. 1000 people aged 15 – 35 years attending a GUM clinic undergo the new test and the current gold standard nucleic acid amplification test (NAAT) to confirm the diagnosis. Of the 1000 people, 250 are diagnosed with chlamydia. Of the patients diagnosed with chlamydia, 240 test positive with the new diagnostic test and of the patients not diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the positive likelihood ratio (LR+) for this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |

4.8

48

0.08

0.48

clinic undergo the new test and the current gold standard nucleic acid amplification test (NAAT) to confirm the diagnosis. Of the 1000 people, 250 are diagnosed with chlamydia. Of the patients diagnosed with chlamydia, 240 test positive with the new diagnostic test and of the patients not diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the positive likelihood ratio (LR+) for this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |

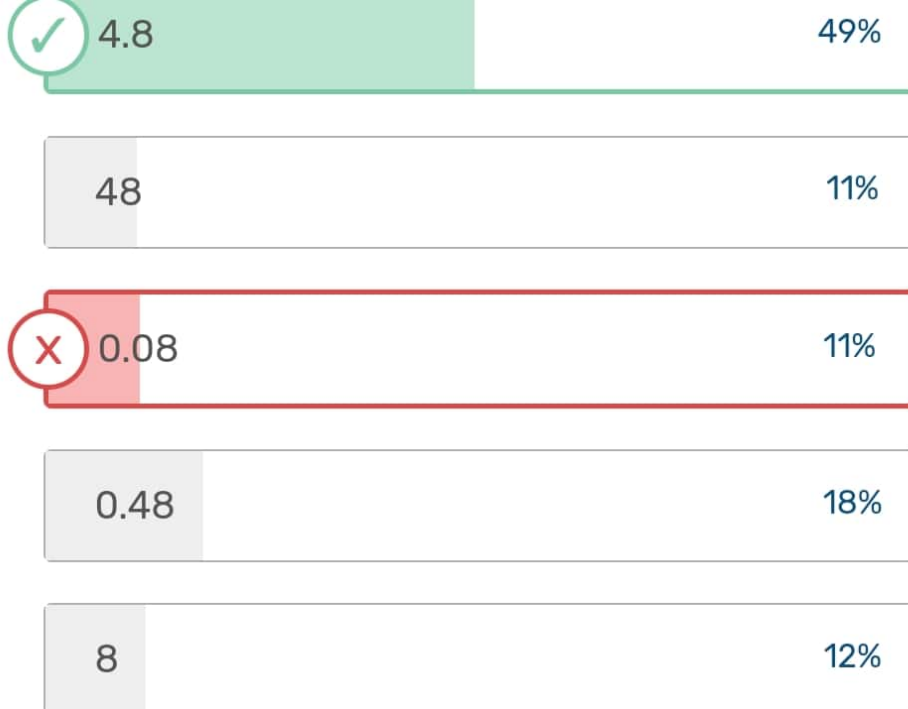
4.8

48

0.08

0.48

8



## ANSWER

- Sensitivity =  $a/(a+c) = 240/250 = 0.96 = 96\%$ 
  - This means that if the patient has chlamydia, there is a 96% chance of the test being positive. The test will only have a 4% false negative rate.
- Specificity =  $d/(b+d) = 600/750 = 0.8 = 80\%$ 
  - This means if the patient does not have chlamydia, there is an 80% chance of the test being negative. The test will have a 20% false positive result.
- Positive likelihood ratio (LR+) =  
 $\text{sensitivity}/(1-\text{specificity}) = 0.96/(1 - 0.8) = 4.8$

0.48

18%

8

12%

## ANSWER

---

- Sensitivity =  $a/(a+c) = 240/250 = 0.96 = 96\%$ 
  - This means that if the patient has chlamydia, there is a 96% chance of the test being positive. The test will only have a 4% false negative rate.
- Specificity =  $d/(b+d) = 600/750 = 0.8 = 80\%$ 
  - This means if the patient does not have chlamydia, there is an 80% chance of the test being negative. The test will have a 20% false positive result.
- Positive likelihood ratio (LR+) =  
 $\text{sensitivity}/(1-\text{specificity}) = 0.96/(1 - 0.8) = 4.8$ 
  - This means that a positive test result is 4.8 times more likely to be seen in a patient who has chlamydia than in a patient who does not have chlamydia.

 QUESTION 68

Regarding the box and whisker plot, which of the following statements is CORRECT:

The rectangle contains 50% of the data values.

The ends of the rectangle correspond to the maximum and minimum value.

The whiskers represent the interquartile range.

Outliers cannot be plotted on a box and whisker plot.

The rectangle represents the range of the data.


Regarding the box and whisker plot, which of the following statements is CORRECT:

- ☒ The rectangle contains 50% of the data values. 52%
- ☐ The ends of the rectangle correspond to the maximum and minimum value. 10%
- ☒ The whiskers represent the interquartile range. 21%
- ☐ Outliers cannot be plotted on a box and whisker plot. 8%
- ☐ The rectangle represents the range of the data. 10%

## ANSWER

A boxplot is a vertical or horizontal rectangle used to display the interquartile range, with the ends of the rectangle corresponding to the upper and lower quartiles of the data values. The box contains 50% of the data values. A line drawn through the rectangle corresponds to the median value. Whiskers, starting at the ends of

The ends of the rectangle correspond to the maximum and minimum value. 10%

 The whiskers represent the interquartile range. 21%

Outliers cannot be plotted on a box and whisker plot. 8%

The rectangle represents the range of the data. 10%

## ANSWER

A boxplot is a vertical or horizontal rectangle used to display the interquartile range, with the ends of the rectangle corresponding to the upper and lower quartiles of the data values. The box contains 50% of the data values. A line drawn through the rectangle corresponds to the median value. Whiskers, starting at the ends of the rectangle usually indicate the minimum and maximum values, therefore the entire box and whisker plot represents the range. Any outliers can be plotted independent of the box and whisker plot.





QUESTION 69

A new anti-tuberculosis drug has been tested against usual therapy, and is found to reduce the risk of death from 30 in 1000 to 10 in 1000. How many patients would require treatment in order to prevent 10 extra deaths from tuberculosis:

25

50

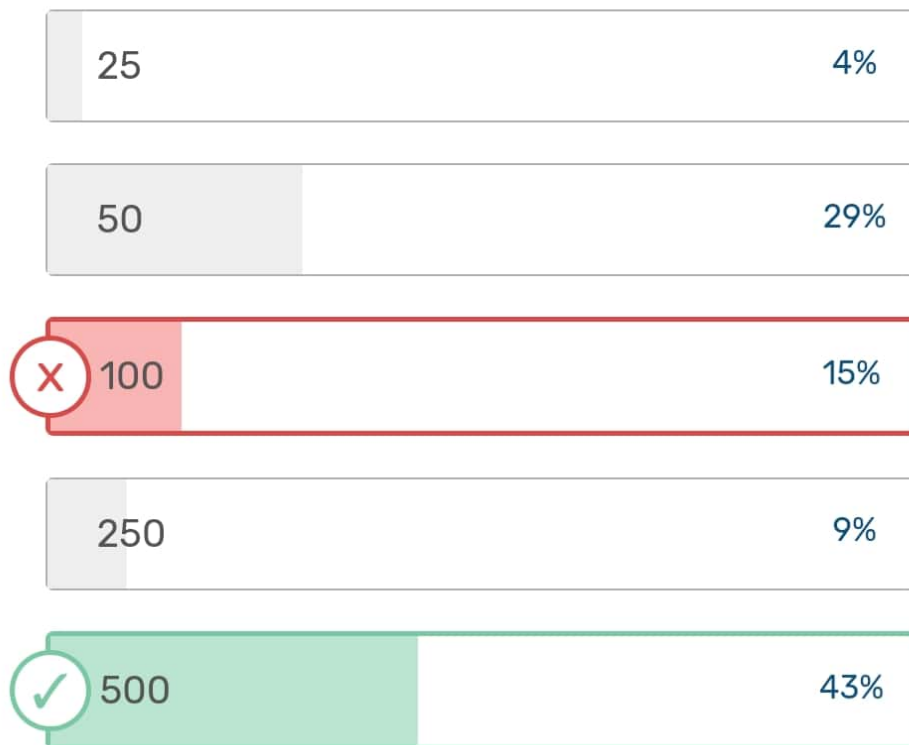
100

250

500



A new anti-tuberculosis drug has been tested against usual therapy, and is found to reduce the risk of death from 30 in 1000 to 10 in 1000. How many patients would require treatment in order to prevent 10 extra deaths from tuberculosis:

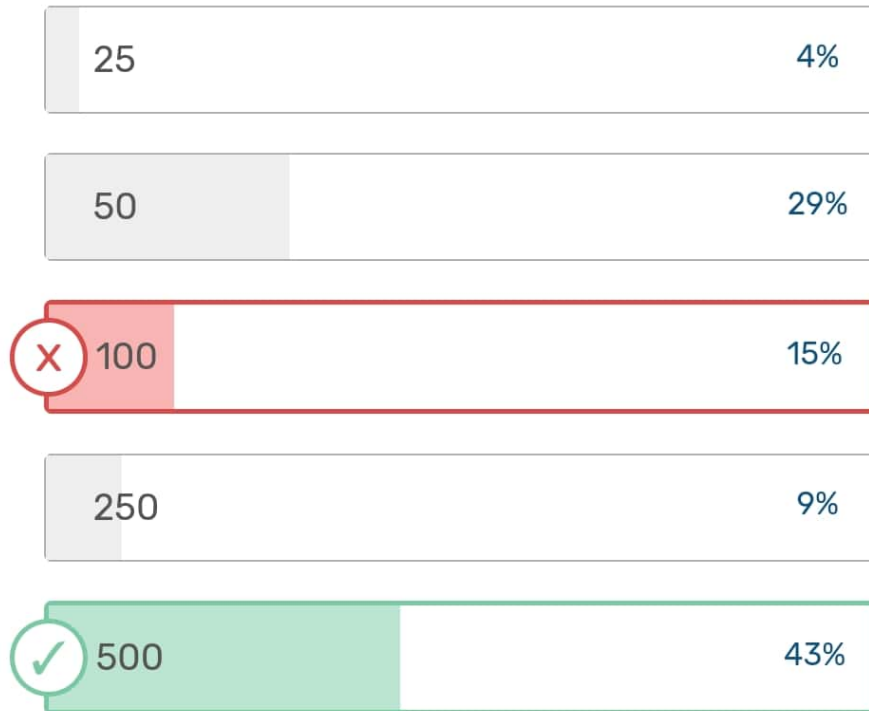


## ANSWER

Absolute risk reduction (ARR) of treatment = risk of death in control group (usual therapy) - risk of death in treatment group

$$\text{ARR} = (30/1000) - (10/1000) = 20/1000 = 0.02$$

$$\text{Number needed to treat (NNT)} = 1/\text{ARR} = 1/0.02 = 50$$



## ANSWER

Absolute risk reduction (ARR) of treatment = risk of death in control group (usual therapy) - risk of death in treatment group

$$\text{ARR} = (30/1000) - (10/1000) = 20/1000 = 0.02$$

$$\text{Number needed to treat (NNT)} = 1/\text{ARR} = 1/0.02 = 50$$

Therefore 50 people would need to be treated to prevent one extra death.

Therefore 500 people would need to be treated to prevent ten extra deaths.





## QUESTION 70

$P = 0.05$  means that the probability of an observed difference between two groups having happened by chance is:

[See Answer](#)



## QUESTION 70

$P = 0.05$  means that the probability of an observed difference between two groups having happened by chance is:

0.05%

14%

0.50%

4%



5%

69%

50%

9%

1%

5%

## ANSWER

The p value gives the probability of observing a difference between the two groups if the null hypothesis is true. In other words, the probability that any observed difference has happened by chance and not by any true difference. a  $p = 0.05$

$P = 0.05$  means that the probability of an observed difference between two groups having happened by chance is:

|                                     |     |
|-------------------------------------|-----|
| 0.05%                               | 14% |
| 0.50%                               | 4%  |
| <input checked="" type="radio"/> 5% | 69% |
| 50%                                 | 9%  |
| 1%                                  | 5%  |

## ANSWER

The p value gives the probability of observing a difference between the two groups if the null hypothesis is true. In other words, the probability that any observed difference has happened by chance and not by any true difference. a  $p = 0.05$  means there is a 5% chance that the observed value has happened by chance.



Save

End Session



A cohort study is performed where researchers observe 100 smokers and 100 non-smokers for 20 years to look for the development of lung cancer. Thirty of the smokers and ten of the non-smokers develop lung cancer within the time period. What is the risk ratio for developing lung cancer in smokers compared to non-smokers:

|            | Lung cancer p |
|------------|---------------|
| Smoker     | 30            |
| Non-Smoker | 10            |
| Total      | 40            |

|            |    |
|------------|----|
| Non-Smoker | 10 |
| Total      | 40 |

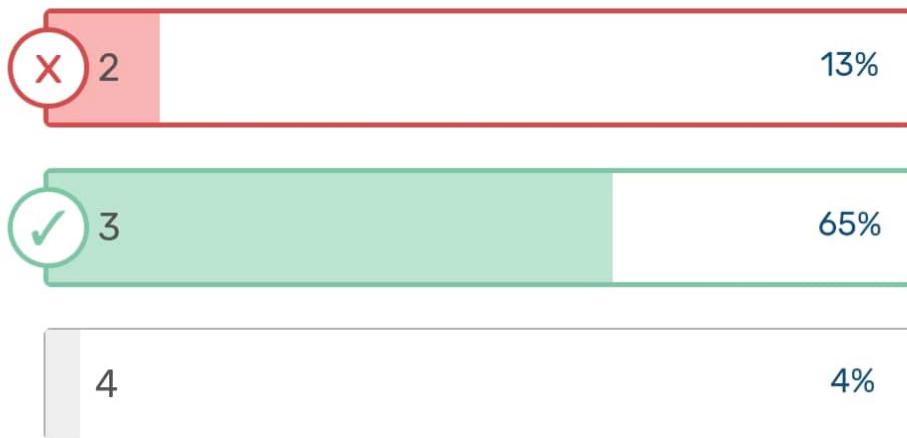


## ANSWER

The risk ratio for developing lung cancer in smokers compared to non-smokers = Risk of lung cancer in smokers/risk of lung cancer in non-smokers

Risk of lung cancer in smokers = number of smokers who developed lung cancer/total number of smokers =  $30/100 = 0.3 = 30\%$

Risk of lung cancer in non-smokers = number of non-smokers who developed lung cancer/total number of non-smokers =  $10/100 = 0.1 = 10\%$



## ANSWER

The risk ratio for developing lung cancer in smokers compared to non-smokers = Risk of lung cancer in smokers/risk of lung cancer in non-smokers

Risk of lung cancer in smokers = number of smokers who developed lung cancer/total number of smokers =  $30/100 = 0.3 = 30\%$

Risk of lung cancer in non-smokers = number of non-smokers who developed lung cancer/total number of non-smokers =  $10/100 = 0.1 = 10\%$

Risk Ratio =  $0.3/0.1 = 3$ .

Therefore smokers in this cohort were 3 times more likely to develop lung cancer than non-smokers.





## QUESTION 72

Which of the following terms describes the proportion of patients without a disease who are correctly identified by a test as not having the disease:

Specificity

Sensitivity

Negative predictive value

Positive predictive value

Negative likelihood ratio



See Answer



## QUESTION 72

Which of the following terms describes the proportion of patients without a disease who are correctly identified by a test as not having the disease:



Specificity

62%

Sensitivity

9%



Negative predictive value

24%

Positive predictive value

3%

Negative likelihood ratio

3%

## ANSWER

Specificity is the proportion of patients without the disease (true negatives) who are correctly identified by the test as not having the disease = true negative rate.



Which of the following terms describes the proportion of patients without a disease who are correctly identified by a test as not having the disease:

- ☒ Specificity 62%
- ☐ Sensitivity 9%
- ☒ Negative predictive value 24%
- ☐ Positive predictive value 3%
- ☐ Negative likelihood ratio 3%

## ANSWER

Specificity is the proportion of patients without the disease (true negatives) who are correctly identified by the test as not having the disease = true negative rate.



Save

End Session



Regarding correlation, which of the following statements is CORRECT:

The Pearson correlation coefficient is used for non-normally distributed data.

Positive correlation indicates a cause and effect relationship.

The correlation coefficient is normally denoted by 'alpha'.

A positive correlation coefficient means that there is strong correlation between two variables.


The closer that the correlation is to 1, the closer the points are to a straight line.




The Pearson correlation coefficient is used for non-normally distributed data. 5%

Positive correlation indicates a cause and effect relationship. 10%


The correlation coefficient is normally denoted by 'alpha'. 6%


 A positive correlation coefficient means that there is strong correlation between two variables. 40%

 The closer that the correlation is to 1, the closer the points are to a straight line. 39%

## ANSWER

If there is a perfect relationship between the two variables, then  $r = 1$  (+ or -). The closer that  $r$  is to 1, the greater the strength of correlation (and the closer the points are to a straight line). A positive correlation coefficient means that as the value of one variable increases, the value of the other variable increases. The correlation coefficient is normally denoted by  $r$ . Correlation does not give information about a cause and effect.

 A positive correlation coefficient means that there is strong correlation between two variables. 40%

 The closer that the correlation is to 1, the closer the points are to a straight line. 39%

## ANSWER

If there is a perfect relationship between the two variables, then  $r = 1$  (+ or -). The closer that  $r$  is to 1, the greater the strength of correlation (and the closer the points are to a straight line). A positive correlation coefficient means that as the value of one variable increases, the value of the other variable increases. The correlation coefficient is normally denoted by  $r$ . Correlation does not give information about a cause and effect relationship. The Pearson's correlation coefficient can be used if the values are sampled from normal populations, otherwise the non-parametric equivalent, the Spearman's rank correlation coefficient, can be used.



Save

End Session



## QUESTION 74

Which of the following is an example of a parametric test:

Mann-Whitney U-test

Wilcoxon matched pairs test

Kruskal-Wallis test

Friedman's test

Analysis of variance (ANOVA) test



See Answer

Save

End Session



## QUESTION 74

Which of the following is an example of a parametric test:

Mann-Whitney U-test

11%

Wilcoxon matched pairs test

6%



Kruskal-Wallis test

5%

Friedman's test

10%



Analysis of variance (ANOVA) test

68%

### ANSWER

The analysis of variance (ANOVA) test is a parametric test.





The results from the study investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the specificity calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

$$a/(a+c)$$

$$d/(c+d)$$

$$a/(a+b)$$

$$(a+c)/a$$

investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the specificity calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

✓

$d/(b+d)$

69%

$a/(a+c)$

10%

✗

$d/(c+d)$

14%

$a/(a+b)$

6%

$(a+c)/a$

2%

## ANSWER

Specificity is the proportion of patients without the disease (true negatives) who are correctly identified by the test as not having the disease = true negative rate.



Which of the following is NOT a disadvantage of a case-control study used to identify past exposure to a risk factor in patients with a disease:

Prone to confounding

Unsuitable for rare risk factors

Difficult to establish temporal relationship between exposure and outcome

Cannot be used to quantify absolute risk

Subject to loss to follow up bias



See Answer

Save

End Session



Which of the following is NOT a disadvantage of a case-control study used to identify past exposure to a risk factor in patients with a disease:

Prone to confounding

6%

Unsuitable for rare risk factors

22%



Difficult to establish temporal relationship between exposure and outcome

15%

Cannot be used to quantify absolute risk

14%



Subject to loss to follow up bias

45%

## ANSWER

Advantages:

- relatively quick
- relatively cheap and easy to perform
- particularly suitable for studying associations between an exposure and an outcome when the outcome is uncommon or if the outcome occurs decades after



Subject to loss to follow up bias

43%

## ANSWER

### Advantages:

- relatively quick
- relatively cheap and easy to perform
- particularly suitable for studying associations between an exposure and an outcome when the outcome is uncommon or if the outcome occurs decades after exposure
- a wide range of risk factors can be investigated in each study

### Disadvantages:

- subject to recall bias
- unlike in a whole population study, absolute risk cannot be quantified
- temporal relationship between exposure and outcome can be difficult to establish
- unsuitable for rare risk factors
- prone to confounding



Save

End Session



SCORE

13%



Which of the following best describes the positive predictive value of a diagnostic test:

The proportion of individuals with a positive test result who have the disease

The ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease

The proportion of patients with the disease who are correctly identified by the test as having the disease

The proportion of individuals with a negative test result who do not have the disease

The proportion of patients without the disease who are correctly identified by the test as not having the disease

The proportion of individuals with a positive test result who have the disease

The ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease

The proportion of patients with the disease who are correctly identified by the test as having the disease

The proportion of individuals with a negative test result who do not have the disease

The proportion of patients without the disease who are correctly identified by the test as not having the disease



See Answer

Save

End Session

✓ The proportion of individuals with a positive test result who have the disease 57%

The ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease 15%

✗ The proportion of patients with the disease who are correctly identified by the test as having the disease 25%

The proportion of individuals with a negative test result who do not have the disease 1%

The proportion of patients without the disease who are correctly identified by the test as not having the disease 3%

## ANSWER

Positive predictive value (PPV) is the proportion of individuals with a positive test result who actually have the disease.





As part of a quality improvement project you are handing out patient satisfaction questionnaires to patients who have presented with low back pain. Patients are asked to describe how satisfied they were with their care as:

- Very satisfied
- Somewhat satisfied
- Neither satisfied nor dissatisfied
- Somewhat dissatisfied
- Very dissatisfied

What type of data is being collected:

Quantitative data

Discrete data

Continuous data

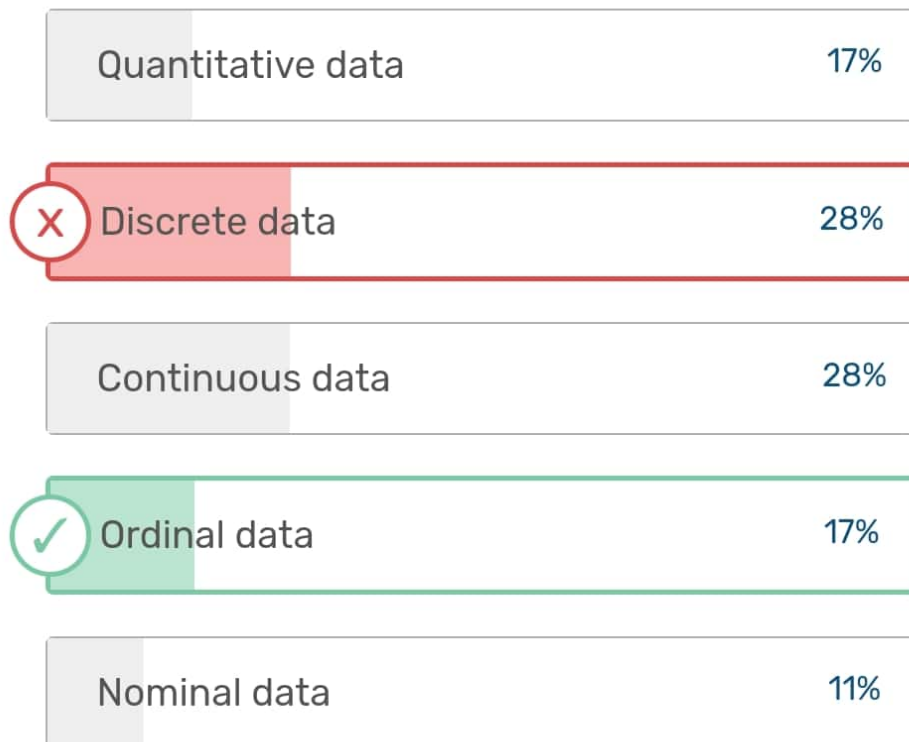
Ordinal data

Nominal data

satisfaction questionnaires to patients who have presented with low back pain. Patients are asked to describe how satisfied they were with their care as:

- Very satisfied
- Somewhat satisfied
- Neither satisfied nor dissatisfied
- Somewhat dissatisfied
- Very dissatisfied

What type of data is being collected:



## ANSWER

Ordinal data is categorical data that is ordered in some way e.g. disease staging system, pain scoring system.



## QUESTION 79

What is the mode of the following data set: 13, 18, 13, 14, 13, 16, 14, 21, 13



See Answer

Save

End Session



## QUESTION 79

What is the mode of the following data set: 13, 18, 13, 14, 13, 16, 14, 21, 13

|  |    |     |
|--|----|-----|
|  | 13 | 90% |
|  | 14 | 3%  |
|  | 16 | 3%  |
|  | 18 | 2%  |
|  | 21 | 2%  |

## ANSWER

The mode is the most common occurring number = 13.





## QUESTION 80

Which of the following is an example of nominal data:

Disease staging system

Number of children

Height

Marital status

Body temperature



See Answer

Save

End Session



## QUESTION 80

Which of the following is an example of nominal data:

Disease staging system

13%

Number of children

16%



Height

7%



Marital status

55%

Body temperature

9%

### ANSWER

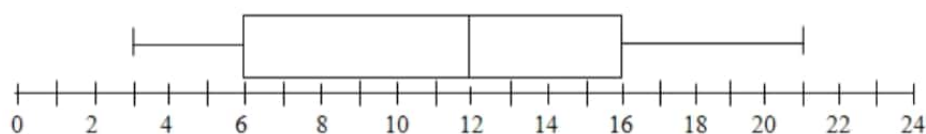
Nominal data is data where the categories have names but are not ordered in any way e.g. blood group, marital status, gender.





## QUESTION 81

Regarding the following box and whisker plot, what is the interquartile range of the data:



6

10

12

18

21





## QUESTION 81

Regarding the following box and whisker plot, what is the interquartile range of the data:



6

6%



10

61%



12

18%

18

13%

21

2%

## ANSWER

The interquartile range is represented by the rectangle. The interquartile range is the difference between the upper quartile and the lower quartile, demarcated by the ends of the rectangle. The interquartile range =  $16 - 6 = 10$ .



In hypothesis testing, what does beta denote:

The chance of making a type I error

The chance of making a type II error

The chance of detecting a statistically significant difference where one is present

The chance of correctly accepting the null hypothesis when it is true

The chance of correctly rejecting the null hypothesis when it is false



See Answer

Save

End Session



In hypothesis testing, what does beta denote:

The chance of making a type I error

4%



The chance of making a type II error

69%

The chance of detecting a statistically significant difference where one is present

15%



The chance of correctly accepting the null hypothesis when it is true

7%

The chance of correctly rejecting the null hypothesis when it is false

6%

## ANSWER

The probability of making a type II error is denoted by beta ( $\beta$ ).



Save

End Session



An FY2 in ED is collecting retrospective data for an audit on headache. He records the number of patients with headaches seen per day and for each patient who presented to ED with a headache he collects data on the length of time waiting to see the triage nurse, the pain score at triage (Faces Pain Rating Scale) and pain relief given in the ED (e.g. simple non-opioid analgesia, opioid analgesia and nil).

The length of time waiting to be seen by triage is an example of which of the following types of data:

Categorical

Nominal

Ordinal

Discrete

Continuous

patient who presented to ED with a headache he collects data on the length of time waiting to see the triage nurse, the pain score at triage (Faces Pain Rating Scale) and pain relief given in the ED (e.g. simple non-opioid analgesia, opioid analgesia and nil). The length of time waiting to be seen by triage is an example of which of the following types of data:

|                                  |             |     |
|----------------------------------|-------------|-----|
| <input type="radio"/>            | Categorical | 7%  |
| <input type="radio"/>            | Nominal     | 11% |
| <input checked="" type="radio"/> | Ordinal     | 12% |
| <input type="radio"/>            | Discrete    | 15% |
| <input checked="" type="radio"/> | Continuous  | 56% |

## ANSWER

This is an example of continuous quantitative data where there is no limitation on the value the measured data can take.





Regarding linear relationships between two variables, what does a positive correlation coefficient of +1 indicate:

The two variables are inversely proportional

As one variable increases, the other variable decreases

As one variable decreases, the other variable increases

The two variables are directly proportional

There is low correlation between the two variables




See Answer




Regarding linear relationships between two variables, what does a positive correlation coefficient of +1 indicate:

The two variables are inversely proportional 3%

As one variable increases, the other variable decreases 9%

 As one variable decreases, the other variable increases 5%

 The two variables are directly proportional 61%

There is low correlation between the two variables 22%

## ANSWER

A positive correlation coefficient means that the two variables are directly proportional e.g. height and weight in healthy growing children.





Regarding forest plots in meta-analysis, the horizontal lines extending from each square represent:

The 95% confidence interval for the effect size estimate of a study.

The effect size estimate of a study.

The weight given to a study.

The 95% confidence interval for the summary effect size estimate.

The summary effect size estimate.



See Answer

Save

End Session



Regarding forest plots in meta-analysis, the horizontal lines extending from each square represent:



The 95% confidence interval for the effect size estimate of a study.

48%

The effect size estimate of a study.

12%

The weight given to a study.

14%



The 95% confidence interval for the summary effect size estimate.

19%

The summary effect size estimate.

7%

## ANSWER

The right-hand column is a plot of the measure of effect size, for example an odds ratio (OR), for each of these studies, represented by a square:

- The midpoint of the square represents the effect size.
- The width of the horizontal lines extending from the square represents the 95% confidence interval for this effect size.
- The area of each square is proportional to

|   |  |     |
|---|--|-----|
| ✓ | The 95% confidence interval for the effect size estimate of a study. | 48% |
|   | The effect size estimate of a study.                                 | 12% |
|   | The weight given to a study.   | 14% |
| ✗ | The 95% confidence interval for the summary effect size estimate.    | 19% |
|   | The summary effect size estimate.                                    | 7%  |

## ANSWER

The right-hand column is a plot of the measure of effect size, for example an odds ratio (OR), for each of these studies, represented by a square:

- The midpoint of the square represents the effect size.
- The width of the horizontal lines extending from the square represents the 95% confidence interval for this effect size.
- The area of each square is proportional to the weight given to that study in the meta-analysis (studies with larger sample sizes, and more effect sizes are given more weight).

A new test is being developed to diagnose chlamydia. 1000 people aged 15 – 35 years attending a GUM clinic undergo the new test and the current gold standard nucleic acid amplification test (NAAT) to confirm the diagnosis. Of the 1000 people, 250 are diagnosed with chlamydia. Of the patients diagnosed with chlamydia, 240 test positive with the new diagnostic test and of the patients not diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the negative predictive value of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |

62%

75%

80%

96%

current gold standard nucleic acid amplification test (NAAT) to confirm the diagnosis. Of the 1000 people, 250 are diagnosed with chlamydia. Of the patients diagnosed with chlamydia, 240 test positive with the new diagnostic test and of the patients not diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the negative predictive value of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a = 240       |
| Negative test | c = 10        |
| Total         | 250           |

62%

75%

80%

96%

98%

diagnosed with chlamydia, 100 test positive with the new diagnostic test. What is the negative predictive value of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |

62%

12%

75%

6%

X

80%

16%

96%

6%

✓

98%

61%

## ANSWER

- Negative predictive value (NPV) =  $d/(c+d) = 600/610 = 0.98 = 98\%$ 
  - This means there is a 98% chance, if the test is negative, that the patient does not have chlamydia.

|               |               |
|---------------|---------------|
|               | Chlamydia Yes |
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |



## ANSWER

- Negative predictive value (NPV) =  $d/(c+d) = 600/610 = 0.98 = 98\%$ 
  - This means there is a 98% chance, if the test is negative, that the patient does not have chlamydia.



Save

End Session



Regarding statistical error, which of the following statements is INCORRECT:

Conventionally, alpha is taken as 0.05.

A type II error occurs when the null hypothesis is accepted when it is false.

The chance of making a type II error is denoted as beta.

The power of a study is the probability that a study will detect a statistically significant difference where one is present.

A narrow confidence interval suggests a low powered study.







Regarding statistical error, which of the following statements is INCORRECT:

Conventionally, alpha is taken as 0.05. 8%

A type II error occurs when the null hypothesis is accepted when it is false. 12%

The chance of making a type II error is denoted as beta. 7%

 The power of a study is the probability that a study will detect a statistically significant difference where one is present. 24%

 A narrow confidence interval suggests a low powered study. 49%

## ANSWER

Inspection of the confidence intervals for the effect of interest gives an indication of whether the power of the test was adequate; a wide confidence interval suggests low power. The most common reason for type II error is that the

A type II error occurs when the null hypothesis is accepted when it is false. 12%

The chance of making a type II error is denoted as beta. 7%

The power of a study is the probability that a study will detect a statistically significant difference where one is present. 24%

A narrow confidence interval suggests a low powered study. 49%

## ANSWER

Inspection of the confidence intervals for the effect of interest gives an indication of whether the power of the test was adequate; a wide confidence interval suggests low power. The most common reason for type II error is that the study is too small.



Save

End Session





A group of patients enrolling for a trial had a normal distribution for weight. The mean weight of the patients was 80 kg. The standard deviation was calculated to be 5 kg. Regarding this group of patients, which of the following statements is most accurate:

95% of the patients will weigh between 75 and 85 kg.

68% of the patients will weigh between 75 and 85 kg.

99% of the patients will weigh between 70 and 90 kg.

95% of the patients will weigh between 65 and 95 kg.

68% of patients will weigh between 77.5 kg and 82.5 kg



The mean weight of the patients was 80 kg. The standard deviation was calculated to be 5 kg. Regarding this group of patients, which of the following statements is most accurate:

95% of the patients will weigh between 75 and 85 kg. 25%



68% of the patients will weigh between 75 and 85 kg. 57%



99% of the patients will weigh between 70 and 90 kg. 7%

95% of the patients will weigh between 65 and 95 kg. 5%

68% of patients will weigh between 77.5 kg and 82.5 kg 5%

## ANSWER

If we know the mean and standard deviation of a set of normally distributed observations, we can estimate the range of values that would be expected to include certain proportions of observations:

- A range of one SD above and below the mean ( $\mu \pm 1 \text{ SD}$ ) includes 68.2% of the

95% of the patients will weigh  
between 65 and 95 kg.

5%

68% of patients will weigh between  
77.5 kg and 82.5 kg

5%

## ANSWER

If we know the mean and standard deviation of a set of normally distributed observations, we can estimate the range of values that would be expected to include certain proportions of observations:

- A range of one SD above and below the mean ( $\pm 1$  SD) includes 68.2% of the sample values therefore about 68% of the patients will weigh between 75 and 85 kg.
- $\pm 2$  SD includes 95.4% of the values therefore about 95% of the patients will weigh between 70 and 90 kg.
- $\pm 3$  SD includes 99.7% of the values therefore almost 100% of the patients will weigh between 65 and 95 kg.



Save

End Session

 QUESTION 89

Regarding statistical error, which of the following statements is CORRECT:

A type I error occurs when the null hypothesis is accepted when it is false.

The maximum probability of making a type I error accepted is denoted by beta.


The null hypothesis is rejected if the p value is more than alpha.

The power of a study =  $1 - \alpha$ .

For a fixed sample size, the probability of a type I error increases as the probability of a type II error decreases.


Regarding statistical error, which of the following statements is CORRECT:

A type I error occurs when the null hypothesis is accepted when it is false. 18%

 The maximum probability of making a type I error accepted is denoted by beta. 5%

The null hypothesis is rejected if the p value is more than alpha. 15%

The power of a study =  $1 - \alpha$ . 16%

 For a fixed sample size, the probability of a type I error increases as the probability of a type II error decreases. 46%

## ANSWER

The probability of a type I error increases as the probability of a type II error decreases. A type I error occurs when the null hypothesis is wrongly rejected when it is actually true and we conclude that there is a difference of effect when in reality

beta.

The null hypothesis is rejected if the p value is more than alpha. 15%

The power of a study =  $1 - \alpha$ . 16%

✓ For a fixed sample size, the probability of a type I error increases as the probability of a type II error decreases. 46%

## ANSWER

The probability of a type I error increases as the probability of a type II error decreases. A type I error occurs when the null hypothesis is wrongly rejected when it is actually true and we conclude that there is a difference of effect when in reality there is none (a false positive result). The probability of making a type I error is denoted by alpha ( $\alpha$ ), the predetermined significance level of the test (conventionally  $\alpha = 0.05$ ). We reject the null hypothesis is  $p < \alpha$ . A type II error occurs when the null hypothesis is wrongly accepted when it is actually false and we conclude that there is no evidence of a difference in effect when one really exists (a false negative result). The probability of making a type II error is denoted by beta ( $\beta$ ). The power of a study =  $1 - \beta$ .



The results from the study investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the negative predictive value calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

$$a/(a+c)$$

$$d/(c+d)$$

$$a/(a+b)$$

$$(a+c)/a$$



diagnostic test can be displayed in the following format. How is the negative predictive value calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

14%

$$a/(a+c)$$

3%



$$d/(c+d)$$

79%



$$a/(a+b)$$

3%

$$(a+c)/a$$

2%

## ANSWER

Negative predictive value (NPV) is the proportion of individuals with a negative test result who do not have the disease.

$$NPV = d/(c+d)$$



Which of the following is NOT a benefit of systematic review:

Systematic review is quicker and less costly to perform than a new study.

Systematic review improves the reliability and accuracy of recommendations.

Systematic review has greater power to detect effects of interest.

Systematic review allows results to be generalised to a wider population and broader setting.

Systematic review is not affected by bias.




See Answer



Which of the following is NOT a benefit of systematic review:

Systematic review is quicker and less costly to perform than a new study. 9%

Systematic review improves the reliability and accuracy of recommendations. 6%

 Systematic review has greater power to detect effects of interest. 9%

Systematic review allows results to be generalised to a wider population and broader setting. 12%


 Systematic review is not affected by bias. 64%

## ANSWER

Systematic review is not immune to bias. Ensuring all papers have an equal chance of being in the systematic review (e.g. including unpublished studies, studies in other languages etc.) minimises bias.

Systematic review is quicker and less costly to perform than a new study. 9%

Systematic review improves the reliability and accuracy of recommendations. 6%

 Systematic review has greater power to detect effects of interest. 9%

Systematic review allows results to be generalised to a wider population and broader setting. 12%

 Systematic review is not affected by bias. 64%

## ANSWER

Systematic review is not immune to bias. Ensuring all papers have an equal chance of being in the systematic review (e.g. including unpublished studies, studies in other languages etc.) minimises bias.



Save

End Session



## QUESTION 92

Regarding a case-control study, which of the following statements is INCORRECT:

It is retrospective.

It is observational.

The usual outcome measure is the relative risk.

It is particularly suitable for rare diseases.

It is subject to recall bias.





## QUESTION 92

Regarding a case-control study, which of the following statements is INCORRECT:

It is retrospective.

11%

It is observational.

8%



The usual outcome measure is the relative risk.

45%

It is particularly suitable for rare diseases.

25%

It is subject to recall bias.

12%

## ANSWER

A case-control study is a longitudinal, retrospective, observational study which investigates the relationship between a risk factor and one or more outcomes. This is done by selecting patients who already have a specific disease (cases), matching them to patients who do not (controls) and then collecting data from the patients to compare

It is retrospective.

11%

It is observational.

8%



The usual outcome measure is the relative risk.

45%

It is particularly suitable for rare diseases.

25%

It is subject to recall bias.

12%

## ANSWER

A case-control study is a longitudinal, retrospective, observational study which investigates the relationship between a risk factor and one or more outcomes. This is done by selecting patients who already have a specific disease (cases), matching them to patients who do not (controls) and then collecting data from the patients to compare past exposure to a possible risk factor. The usual outcome measure is the odds ratio.



Save

End Session



## QUESTION 93

Which of the following is an example of continuous data:

Disease staging system

Number of children

Height

Number of nurses working on a night shift

Gender



See Answer



## QUESTION 93

Which of the following is an example of continuous data:

Disease staging system

7%

Number of children

11%



Height

72%



Number of nurses working on a night shift

7%

Gender

4%

## ANSWER

Continuous data is data where there is no limitation on the numerical value that the variable can take e.g. weight, height.



QUESTION 94

Which of the following is an example of discrete data:

Disease staging system

Number of children

Height

Marital status

Gender



See Answer

Save

End Session



## QUESTION 94

Which of the following is an example of discrete data:

Disease staging system

12%



Number of children

57%



Height

12%

Marital status

8%

Gender

11%

### ANSWER

Discrete data is quantitative data that can only take whole numerical values e.g. number of children, number of days missed from work.





Which of the following is NOT an advantage of a case-control study used to identify past exposure to a risk factor in patients with a disease:

Can directly measure absolute and relative risk of a disease

Wide range of risk factors can be investigated in each study

Particularly suitable for studying association between an exposure and an outcome when the outcome is delayed after exposure

Particularly suitable for studying association between an exposure and an outcome when the outcome is uncommon

Relative quick, cheap and easy to perform

association between an exposure and  
an outcome when the outcome is  
delayed after exposure

Particularly suitable for studying 18%  
association between an exposure and  
an outcome when the outcome is  
uncommon

Relative quick, cheap and easy to 11%  
perform

## ANSWER

### Advantages:

- relatively quick
- relatively cheap and easy to perform
- particularly suitable for studying associations between an exposure and an outcome when the outcome is uncommon or if the outcome occurs decades after exposure
- a wide range of risk factors can be investigated in each study

### Disadvantages:

- subject to recall bias
- unlike in a whole population study, absolute risk cannot be quantified
- temporal relationship between exposure and outcome can be difficult to establish
- unsuitable for rare risk factors

used to identify past exposure to a risk factor in patients with a disease:

|   |  |     |
|---|--|-----|
| ✓ | Can directly measure absolute and relative risk of a disease   | 43% |
|   | Wide range of risk factors can be investigated in each study   | 10% |
|   | Particularly suitable for studying association between an exposure and an outcome when the outcome is delayed after exposure | 18% |
| ✗ | Particularly suitable for studying association between an exposure and an outcome when the outcome is uncommon               | 18% |
|   | Relative quick, cheap and easy to perform  | 11% |

## ANSWER

Advantages:

- relatively quick
- relatively cheap and easy to perform
- particularly suitable for studying

Relative quick, cheap and easy to perform

11%

## ANSWER

---

### Advantages:

- relatively quick
- relatively cheap and easy to perform
- particularly suitable for studying associations between an exposure and an outcome when the outcome is uncommon or if the outcome occurs decades after exposure
- a wide range of risk factors can be investigated in each study

### Disadvantages:

- subject to recall bias
- unlike in a whole population study, absolute risk cannot be quantified
- temporal relationship between exposure and outcome can be difficult to establish
- unsuitable for rare risk factors
- prone to confounding



Save

End Session



## QUESTION 96

Which of the following is an example of a parametric test:

Mann-Whitney U test

Wilcoxon matched pairs test

Kruskal-Wallis test

Friedman's test

Student t-test



See Answer

Save

End Session



## QUESTION 96

Which of the following is an example of a parametric test:

Mann-Whitney U test

8%



Wilcoxon matched pairs test

10%

Kruskal-Wallis test

5%

Friedman's test

9%



Student t-test

68%

### ANSWER

The student t-test is a parametric test.



Save

End Session



QUESTION 97

Which of the following is NOT a measure of spread of data:

Variance

Standard deviation

Range

Correlation coefficient

Interquartile range



See Answer

Save

End Session



## QUESTION 97

Which of the following is NOT a measure of spread of data:

Variance

12%



Standard deviation

6%

Range

6%



Correlation coefficient

69%

Interquartile range

8%

## ANSWER

The correlation coefficient is used to denote the strength of a linear relationship between two variables measure in a single group. It is not a measure of the spread of data unlike the alternative answers.



Which of the following best describes the negative predictive value of a diagnostic test:

The proportion of individuals with a positive test result who have the disease

The ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease

The proportion of patients with the disease who are correctly identified by the test as having the disease

The proportion of individuals with a negative test result who do not have the disease

The proportion of patients without the disease who are correctly identified by the test as not having the disease

positive test result who have the disease

The ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease 4%

The proportion of patients with the disease who are correctly identified by the test as having the disease 3%

✓ The proportion of individuals with a negative test result who do not have the disease 66%

The proportion of patients without the disease who are correctly identified by the test as not having the disease 26%

## ANSWER

Negative predictive value (NPV) is the proportion of individuals with a negative test result who do not have the disease.



Save

End Session



A randomised controlled trial (RCT) is carried out to study the effects of a new anticoagulant, Anticlot, on stroke incidence in patients with atrial fibrillation (AF). 500 patients with AF are treated with Anticlot and 500 receive usual therapy (control group). One year later the number of patients who have had a stroke is recorded and shown below. What is the absolute risk reduction of treatment in preventing stroke:

|          | Stroke Yes |
|----------|------------|
| Anticlot | 30         |
| Control  | 50         |
| Total    | 80         |

fibrillation (AF). 500 patients with AF are treated with Anticlot and 500 receive usual therapy (control group). One year later the number of patients who have had a stroke is recorded and shown below. What is the absolute risk reduction of treatment in preventing stroke:

|          | Stroke Yes |
|----------|------------|
| Anticlot | 30         |
| Control  | 50         |
| Total    | 80         |

2%

4%

6%

8%

10%

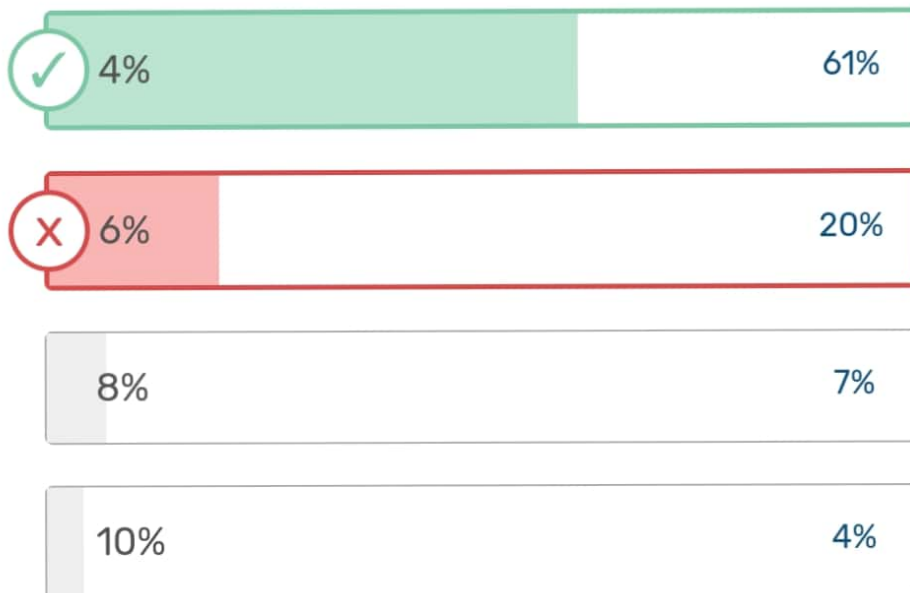


|          |    |
|----------|----|
| Anticlot | 50 |
| Control  | 50 |
| Total    | 80 |



## ANSWER

- The AR of having a stroke in the treatment group (ART) =  $30/500 = 0.06$  (6%)
  - therefore the risk of having a stroke while taking Anticlot is 6%.
- The AR of having a stroke in the control group (ARC) =  $50/500 = 0.1$  (10%)
  - therefore the risk of having a stroke while not taking Anticlot is 10%.
- The ARR =  $ARC - ART = (50/500) - (30/500) = 0.10 - 0.06 = 0.04$  (4%)
  - therefore Anticlot reduces the absolute risk of having a stroke by



## ANSWER

- The AR of having a stroke in the treatment group (ART) =  $30/500 = 0.06$  (6%)
  - therefore the risk of having a stroke while taking Anticlot is 6%.
- The AR of having a stroke in the control group (ARC) =  $50/500 = 0.1$  (10%)
  - therefore the risk of having a stroke while not taking Anticlot is 10%.
- The ARR =  $ARC - ART = (50/500) - (30/500) = 0.10 - 0.06 = 0.04$  (4%)
  - therefore Anticlot reduces the absolute risk of having a stroke by 4%.



Save

End Session



Regarding measures of central tendency, which of the following statements is CORRECT:

The median is not distorted by outliers or skewed data.

The median is the most commonly recurring number.

The median is less than the mean if the data are skewed to the left.

The mean is the middle value of a data set that has been placed in a numerical order.

The mean is the point which has half of the observations above it and half of the observations below it.



See Answer



Regarding measures of central tendency, which of the following statements is CORRECT:

- The median is not distorted by outliers or skewed data.

39%
- The median is the most commonly recurring number.

4%
- The median is less than the mean if the data are skewed to the left.

27%
- The mean is the middle value of a data set that has been placed in a numerical order.

12%
- The mean is the point which has half of the observations above it and half of the observations below it.

19%

## ANSWER

The median is the middle value of a data set that has been placed in numerical order, and the point at which half of the observations lie above it and half below it. The median is not distorted

outliers or skewed data.

The median is the most commonly recurring number.

4%

X

The median is less than the mean if the data are skewed to the left.

27%

The mean is the middle value of a data set that has been placed in a numerical order.

12%

The mean is the point which has half of the observations above it and half of the observations below it.

19%

## ANSWER

The median is the middle value of a data set that has been placed in numerical order, and the point at which half of the observations lie above it and half below it. The median is not distorted by outliers or skewed data. The mode is the most commonly recurring number. The median is similar to the mean if the data is symmetrical, less than the mean if the data are skewed to the right and greater than the mean if the data are skewed to the left.



## QUESTION 101

Regarding standard deviation (SD), which of the following statements is CORRECT:

Standard deviation is typically used for describing the variability of non-normally distributed data.

A data set with a larger standard deviation has a wider spread of data.

The standard deviation is the variance squared.

A range of two SD above and below the mean includes 99% of the sample values.

A range of one SD above and below the mean includes 95% of the sample values.

Regarding standard deviation (SD), which of the following statements is CORRECT:

Standard deviation is typically used for describing the variability of non-normally distributed data. 10%

✓ A data set with a larger standard deviation has a wider spread of data. 52%

✗ The standard deviation is the variance squared. 15%

A range of two SD above and below the mean includes 99% of the sample values. 7%

A range of one SD above and below the mean includes 95% of the sample values. 16%

## ANSWER

The standard deviation is the square root of the variance and provides a measure of the spread of sample values around the sample mean. It is

Standard deviation is typically used for describing the variability of non-normally distributed data. 15%



A data set with a larger standard deviation has a wider spread of data. 52%



The standard deviation is the variance squared. 15%

A range of two SD above and below the mean includes 99% of the sample values. 7%

A range of one SD above and below the mean includes 95% of the sample values. 16%

## ANSWER

The standard deviation is the square root of the variance and provides a measure of the spread of sample values around the sample mean. It is calculated by taking the square root of the average squared difference between each value and the mean and is approximately equivalent to the average difference between the sample value and the mean. A data set with a larger standard deviation has a wider spread of data and vice versa.

A range of one SD above and below the mean includes 95% of the sample values.

16%

## ANSWER

The standard deviation is the square root of the variance and provides a measure of the spread of sample values around the sample mean. It is calculated by taking the square root of the average squared difference between each value and the mean and is approximately equivalent to the average difference between the sample value and the mean. A data set with a larger standard deviation has a wider spread of data and vice versa.

If we know the mean and standard deviation of a set of normally distributed observations, we can estimate the range of values that would be expected to include certain proportions of observations:

- A range of one SD above and below the mean ( $\pm 1$  SD) includes 68.2% of the sample values
- $\pm 2$  SD includes 95.4% of the values
- $\pm 3$  SD includes 99.7% of the values





QUESTION 102

A new chemotherapy drug is being tested. The intervention reduces the risk of death from 10 in 1000 to 5 in 1000. What is the number needed to treat to prevent one death:

15

5

990

200

1000





## QUESTION 102

A new chemotherapy drug is being tested. The intervention reduces the risk of death from 10 in 1000 to 5 in 1000. What is the number needed to treat to prevent one death:

15

3%

5

13%

X

990

12%



200

68%

1000

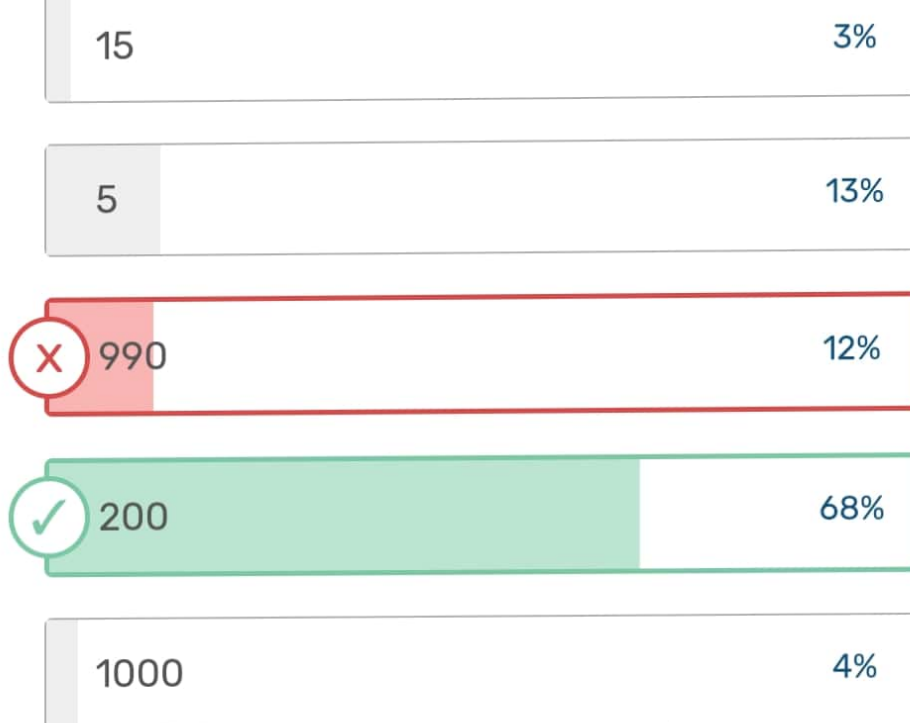
4%

## ANSWER

Absolute risk reduction (ARR) of treatment = risk of death in control group - risk of death in treatment group

$$\text{ARR} = (10/1000) - (5/1000) = 5/1000 = 0.005$$

$$\text{Number needed to treat (NNT)} = 1/\text{ARR} = 1/0.005 = 200$$



## ANSWER

Absolute risk reduction (ARR) of treatment = risk of death in control group - risk of death in treatment group

$$ARR = (10/1000) - (5/1000) = 5/1000 = 0.005$$

$$\text{Number needed to treat (NNT)} = 1/ARR = 1/0.005 = 200$$

Therefore 200 people would need to be treated to prevent one extra death.



Save

End Session



QUESTION 103

What is the interquartile range of the following data set: 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70

15

20

30

35

50



See Answer



## QUESTION 103

What is the interquartile range of the following data set: 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70

15

10%

20

15%



30

42%

35

27%

50

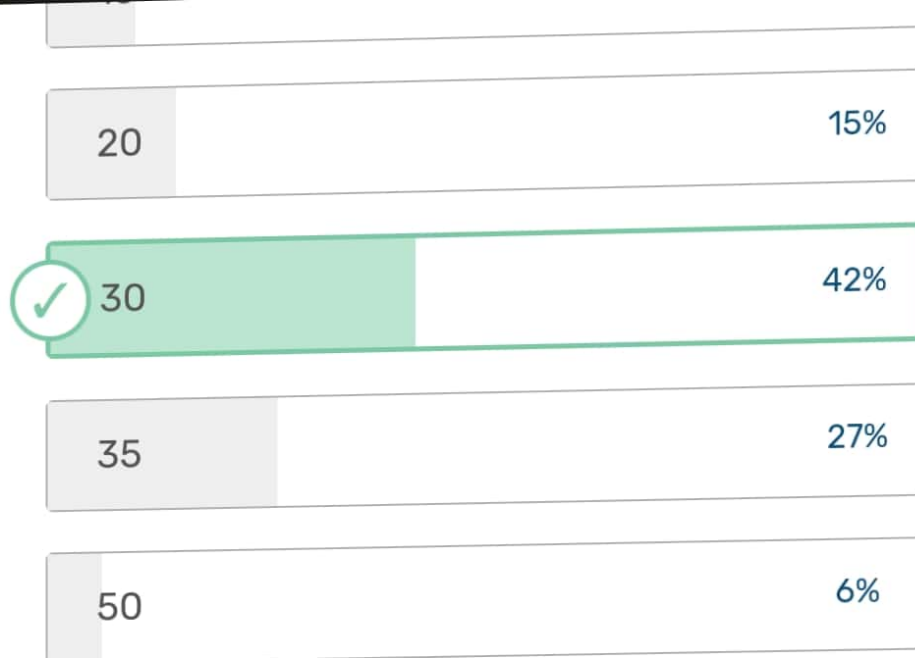
6%

## ANSWER

5, 10, 15 / 20, 25, 30 / 35, 40, 45 / 50, 60, 70

Sample size (n) = 12

Median =  $[(n+1)/2]$ th value =  $(12+1)/2 = 6.5 =$



## ANSWER

5, 10, 15 / 20, 25, 30 / 35, 40, 45 / 50, 60, 70

Sample size (n) = 12

Median =  $[(n+1)/2]$ th value =  $(12+1)/2 = 6.5 =$   
halfway between 6th and 7th value =  $(30 + 35)/2$   
 $= 32.5$

The lower (first) quartile = halfway between 15  
and 20 = 17.5

The upper (third) quartile = halfway between 45  
and 50 = 47.5

The interquartile range is the difference between  
the upper quartile and lower quartile =  $47.5 - 17.5$   
 $= 30$



## QUESTION 104

Which of the following is NOT a disadvantage of a cohort study used to investigate the relationship between exposure to a risk factor and a future outcome:

Can take long periods of time if outcome is delayed

Unsuitable for rare outcomes of interest

Prone to confounding

Subject to recall bias

Subject to loss to follow up bias.





Which of the following is NOT a disadvantage of a cohort study used to investigate the relationship between exposure to a risk factor and a future outcome:

Can take long periods of time if outcome is delayed 12%

Unsuitable for rare outcomes of interest 15%

☒ Prone to confounding 11%

☒ Subject to recall bias 50%

Subject to loss to follow up bias. 13%

## ANSWER

Advantages:

- ideal for studying associations between an exposure and an outcome when the exposure is uncommon
- the time sequence of events can be

Unsuitable for rare outcomes of interest

15%

X

Prone to confounding

11%

✓

Subject to recall bias

50%

Subject to loss to follow up bias.

13%

## ANSWER

### Advantages:

- ideal for studying associations between an exposure and an outcome when the exposure is uncommon
- the time sequence of events can be assessed
- they can provide information on a wide range of disease outcomes
- the absolute and relative risk of disease can be measured directly
- they can give a direct estimation of disease incidence rates

### Disadvantages:

- costly and can take long periods of time if the outcome is delayed
- subject to subject-selection and loss to follow-up bias
- large sample size required for rare outcome of interest so it is not useful for

## ANSWER

---

### Advantages:

- ideal for studying associations between an exposure and an outcome when the exposure is uncommon
- the time sequence of events can be assessed
- they can provide information on a wide range of disease outcomes
- the absolute and relative risk of disease can be measured directly
- they can give a direct estimation of disease incidence rates

### Disadvantages:

- costly and can take long periods of time if the outcome is delayed
- subject to subject-selection and loss to follow-up bias
- large sample size required for rare outcome of interest so it is not useful for rare diseases
- prone to confounding



Save

End Session



## QUESTION 105

Which of the following terms describes the proportion of individuals with a negative test result who actually do not have a disease:

Specificity

Sensitivity

Negative predictive value

Positive predictive value

Negative likelihood ratio



See Answer



## QUESTION 105

Which of the following terms describes the proportion of individuals with a negative test result who actually do not have a disease:

Specificity

28%

Sensitivity

5%



Negative predictive value

61%

Positive predictive value

2%

Negative likelihood ratio

4%

## ANSWER

Negative predictive value (NPV) is the proportion of individuals with a negative test result who do not have the disease.

 QUESTION 106

A range of two standard deviations above and below the mean includes what approximate percentage of the sample values:

[See Answer](#)



## QUESTION 106

A range of two standard deviations above and below the mean includes what approximate percentage of the sample values:

72%

3%

86%

6%



90%

6%



95%

77%

98%

8%

## ANSWER

A range of two SD above and below the mean (+/- 2 SD) includes 95.4% of the sample values.



QUESTION 107

What is the median of the following data set: 13, 18, 13, 14, 13, 16, 14, 21, 13

13

14

15

16

17



See Answer

Save

End Session



## QUESTION 107

What is the median of the following data set: 13, 18, 13, 14, 13, 16, 14, 21, 13

13

15%



14

69%



15

8%

16

6%

17

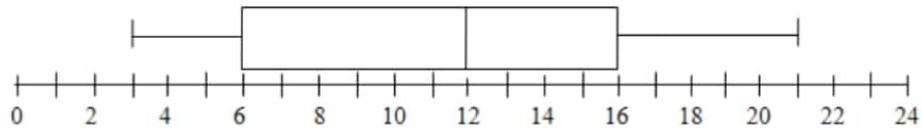
2%

## ANSWER

The median is the middle value of a data set that has been placed in numerical order starting with the smallest value and ending with the largest value. With an even number of data values, the median is the average of the two values that lie on either side of the midline. It is the point which



Regarding the following box and whisker plot, what is the median value of the data:



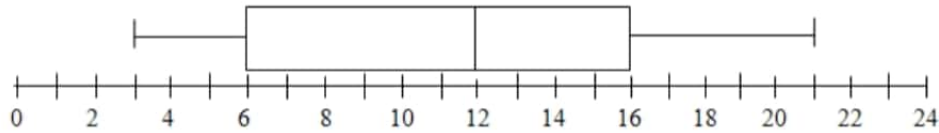
See Answer

Save

End Session



Regarding the following box and whisker plot, what is the median value of the data:



8

1%

10

13%



12

83%

14

1%

16

1%

## ANSWER

The median value is 12, represented by the line drawn through the rectangle.





Regarding the power of a study, which of the following statements is INCORRECT:

The power of a study increases with sample size.

The power of a study is not affected by data variability.

The power of a study is greater for a larger expected effect size.

The power of a study is increased with a larger significance level.

A study power set at 80% accepts the likelihood of a 1 in 5 chance of missing a statistically significant difference where one exists.



Regarding the power of a study, which of the following statements is INCORRECT:

The power of a study increases with sample size. 6%

✓ The power of a study is not affected by data variability. 51%

The power of a study is greater for a larger expected effect size. 9%

✗ The power of a study is increased with a larger significance level. 15%

A study power set at 80% accepts the likelihood of a 1 in 5 chance of missing a statistically significant difference where one exists. 19%

## ANSWER

A study should only be undertaken if the power is at least 80%; a study power set at 80% accepts a likelihood of 1 in 5 (20%) of missing a statistically significant difference where one exists.

The determinants of power are:

^ a larger significance level.

A study power set at 80% accepts the likelihood of a 1 in 5 chance of missing a statistically significant difference where one exists. 19%

## ANSWER

A study should only be undertaken if the power is at least 80%; a study power set at 80% accepts a likelihood of 1 in 5 (20%) of missing a statistically significant difference where one exists.

The determinants of power are:

- the sample size (the power increases with sample size)
- the variability of the observations (the power increases as the variability decreases)
- the effect size of interest (the power is greater for a larger expected effect size)
- and the significance level,  $\alpha$  (the power is greater if the significance level is larger); therefore the probability of a type I error increases as the probability of a type II error decreases.





Regarding standard error of the mean (SEM), which of the following statements is CORRECT:

The SEM for a data sample =  
Standard deviation/Sample size

A large SEM indicates the estimate is precise.

The SEM indicates how much a set of sample values is spread around the sample mean.

The SEM is a measure of the precision of the sample mean as an estimate of the true population mean.

The SEM is increased with increasing sample size.




statements is CORRECT:

The SEM for a data sample = Standard deviation/Sample size 12%

A large SEM indicates the estimate is precise. 4%

The SEM indicates how much a set of sample values is spread around the sample mean. 23%

 The SEM is a measure of the precision of the sample mean as an estimate of the true population mean. 50%

The SEM is increased with increasing sample size. 12%

## ANSWER

The standard error of the mean (SEM) is a measure of the expected spread of sample means (i.e. how much the mean varies on repeated sampling) and gives us a measure of the precision of the sample mean as an estimate of the true population mean.

The SEM for a data sample =  $\sigma/\sqrt{n}$ , where  $\sigma$  =



of the sample mean as an estimate of the true population mean.

The SEM is increased with increasing sample size.

12%

## ANSWER

The standard error of the mean (SEM) is a measure of the expected spread of sample means (i.e. how much the mean varies on repeated sampling) and gives us a measure of the precision of the sample mean as an estimate of the true population mean.

The SEM for a data sample =  $\sigma/\sqrt{n}$ , where  $\sigma$  = standard deviation and  $n$  = sample size.

A large SEM indicates the estimate is imprecise and small SEM indicates that the estimate is precise. The SEM is reduced if the size of the sample is increased or the data is less variable.



Save

End Session





The SEM is increased with increasing sample size. 12%

## ANSWER

The standard error of the mean (SEM) is a measure of the expected spread of sample means (i.e. how much the mean varies on repeated sampling) and gives us a measure of the precision of the sample mean as an estimate of the true population mean.

The SEM for a data sample =  $\sigma/\sqrt{n}$ , where  $\sigma$  = standard deviation and  $n$  = sample size.

A large SEM indicates the estimate is imprecise and small SEM indicates that the estimate is precise. The SEM is reduced if the size of the sample is increased or the data is less variable.



Save

End Session



Regarding the interpretation of risk of an outcome event when comparing a treatment and control group, how is the number needed to treat calculated:

Relative risk in the treatment group/Absolute risk in the control group

1 / Relative risk reduction

1 / Absolute risk reduction

1 / Absolute risk in the treatment group


1 - Absolute risk reduction




See Answer

Regarding the interpretation of risk of an outcome event when comparing a treatment and control group, how is the number needed to treat calculated:

Relative risk in the treatment group/Absolute risk in the control group 10%

 1 / Relative risk reduction 13%

 1 / Absolute risk reduction 54%

1 / Absolute risk in the treatment group 6%

1 - Absolute risk reduction 18%

## ANSWER

The number needed to treat (NNT) is the number of patients who need to be treated with the intervention, compared with the control, in order for one extra patient to experience a beneficial effect.

$$\text{NNT} = 1/\text{absolute risk reduction (ARR)}$$



QUESTION 112

Which of the following does NOT increase the power of a study:

An increase in sample size

A decrease in variability of observations

A larger expected effect of interest

A higher significance level

A lower significance level



See Answer

A decrease in variability of observations 9%

 A larger expected effect of interest 9%

A higher significance level 15%

 A lower significance level 61%

## ANSWER

The determinants of power are:

- the sample size (the power increases with sample size)
- the variability of the observations (the power increases as the variability decreases)
- the effect size of interest (the power is greater for a larger expected effect size)
- and the significance level,  $\alpha$  (the power is greater if the significance level is larger); therefore the probability of a type I error increases as the probability of a type II error decreases.



Save

End Session



## QUESTION 113

Regarding relationships between two variables, what does a negative correlation coefficient of  $-1$  indicate:

The two variables are inversely proportional

There is no correlation between the two variables

As one variable decreases, the other variable decreases

The two variables are directly proportional

Correlation is poor between the two variables



Regarding relationships between two variables, what does a negative correlation coefficient of  $-1$  indicate:

- ☒ The two variables are inversely proportional 67%
- ☐ There is no correlation between the two variables 10%
- ☐ As one variable decreases, the other variable decreases 9%
- ☐ The two variables are directly proportional 5%
- ☐ Correlation is poor between the two variables 9%

## ANSWER

A negative correlation coefficient means that the two variables are inversely proportional e.g. socio-economic class and mortality.



## QUESTION 114

Regarding forest plots in meta-analysis, the midpoint of the square represents:

The 95% confidence interval for the effect size estimate of a study.

The effect size estimate of a study.

The weight given to a study.

The 95% confidence interval for the summary effect size estimate.

The summary effect size estimate.





Regarding forest plots in meta-analysis, the midpoint of the square represents:

The 95% confidence interval for the effect size estimate of a study. 13%

✓ The effect size estimate of a study. 42%

✗ The weight given to a study. 12%

The 95% confidence interval for the summary effect size estimate. 14%

The summary effect size estimate. 19%

## ANSWER

The right-hand column is a plot of the measure of effect size, for example an odds ratio (OR), for each of these studies, represented by a square:

- The midpoint of the square represents the effect size.
- The width of the horizontal lines extending from the square represents the 95% confidence interval for this effect size.
- The area of each square is proportional to

✓ The effect size estimate of a study. 42%

✗ The weight given to a study. 12%

The 95% confidence interval for the summary effect size estimate. 14%

The summary effect size estimate. 19%

## ANSWER

The right-hand column is a plot of the measure of effect size, for example an odds ratio (OR), for each of these studies, represented by a square:

- The midpoint of the square represents the effect size.
- The width of the horizontal lines extending from the square represents the 95% confidence interval for this effect size.
- The area of each square is proportional to the weight given to that study in the meta-analysis (studies with larger sample sizes, and more effect sizes are given more weight).



Save

End Session



## QUESTION 115

Regarding the median, which of the following statements is INCORRECT:

The median is the point which has half of the data values above it, and half of the data values below it.

The median is typically used to compare outcomes in normally distributed data.

The median is not affected by skewed data.

The median ignores most of the data values.

The median is less than the mean if data are skewed to the right.

Regarding the median, which of the following statements is INCORRECT:

The median is the point which has half of the data values above it, and half of the data values below it. 7%



The median is typically used to compare outcomes in normally distributed data. 30%



The median is not affected by skewed data. 31%

The median ignores most of the data values. 12%

The median is less than the mean if data are skewed to the right. 20%

## ANSWER

The median is the middle value of a data set that has been placed in numerical order starting with the smallest value and ending with the largest value. With an even number of data values, the median is the average of the two values that lie on either side of the midline. It is the point which has half of the observations above it and half of

- ✓ The median is typically used to compare outcomes in normally distributed data. 30%
- ✗ The median is not affected by skewed data. 31%
- The median ignores most of the data values. 12%
- The median is less than the mean if data are skewed to the right. 20%

## ANSWER

The median is the middle value of a data set that has been placed in numerical order starting with the smallest value and ending with the largest value. With an even number of data values, the median is the average of the two values that lie on either side of the midline. It is the point which has half of the observations above it and half of the observations below it.

The median is similar to the mean if the data is symmetrical, less than the mean if the data are skewed to the right and greater than the mean if the data are skewed to the left.

The median is not distorted by outliers or skewed

The median ignores most of the data values.

12%

The median is less than the mean if data are skewed to the right.

20%

## ANSWER

The median is the middle value of a data set that has been placed in numerical order starting with the smallest value and ending with the largest value. With an even number of data values, the median is the average of the two values that lie on either side of the midline. It is the point which has half of the observations above it and half of the observations below it.

The median is similar to the mean if the data is symmetrical, less than the mean if the data are skewed to the right and greater than the mean if the data are skewed to the left.

The median is not distorted by outliers or skewed data but ignores most of the data values and so it can be shown to be less efficient than the mean.



Save

End Session



## QUESTION 116

Regarding linear relationships between two variables, what does the correlation coefficient  $r = 0$  indicate:

The two variables are inversely proportional

The two variables are directly proportional

There is no correlation between two variables

There is perfect correlation between two variables

The correlation is not statistically significant

Regarding linear relationships between two variables, what does the correlation coefficient  $r = 0$  indicate:

- The two variables are inversely proportional 4%
- The two variables are directly proportional 6%
- ☒ There is no correlation between two variables 74%
- There is perfect correlation between two variables 7%
- The correlation is not statistically significant 9%

## ANSWER

If there is no correlation between two variables, then  $r = 0$ . The closer that  $r$  is to 0, the weaker the correlation.





## QUESTION 117

Which of the following terms describes the proportion of individuals with a positive test result who actually have a disease:

Specificity

Sensitivity

Negative predictive value

Positive predictive value

Positive likelihood ratio



See Answer

Which of the following terms describes the proportion of individuals with a positive test result who actually have a disease:

Specificity

8%

Sensitivity

31%



Negative predictive value

2%



Positive predictive value

57%

Positive likelihood ratio

2%

## ANSWER

Positive predictive value (PPV) is the proportion of individuals with a positive test result who actually have the disease.



Save

End Session



A randomised controlled trial (RCT) is performed where 1000 men are treated with a lipid-lowering drug, Superstatin, and 1000 given a placebo. Two years later the number of patients who have had a myocardial infarction (MI) is recorded and shown below. What is the relative risk of having an MI in the treatment group compared to the control group:

|             | Myocardial infarction<br>Yes |
|-------------|------------------------------|
| Superstatin | 20                           |
| Placebo     | 40                           |
| Total       | 60                           |

placebo. Two years later the number of patients who have had a myocardial infarction (MI) is recorded and shown below. What is the relative risk of having an MI in the treatment group compared to the control group:

|             | Myocardial infarction<br>Yes |
|-------------|------------------------------|
| Superstatin | 20                           |
| Placebo     | 40                           |
| Total       | 60                           |

0.5%

2%

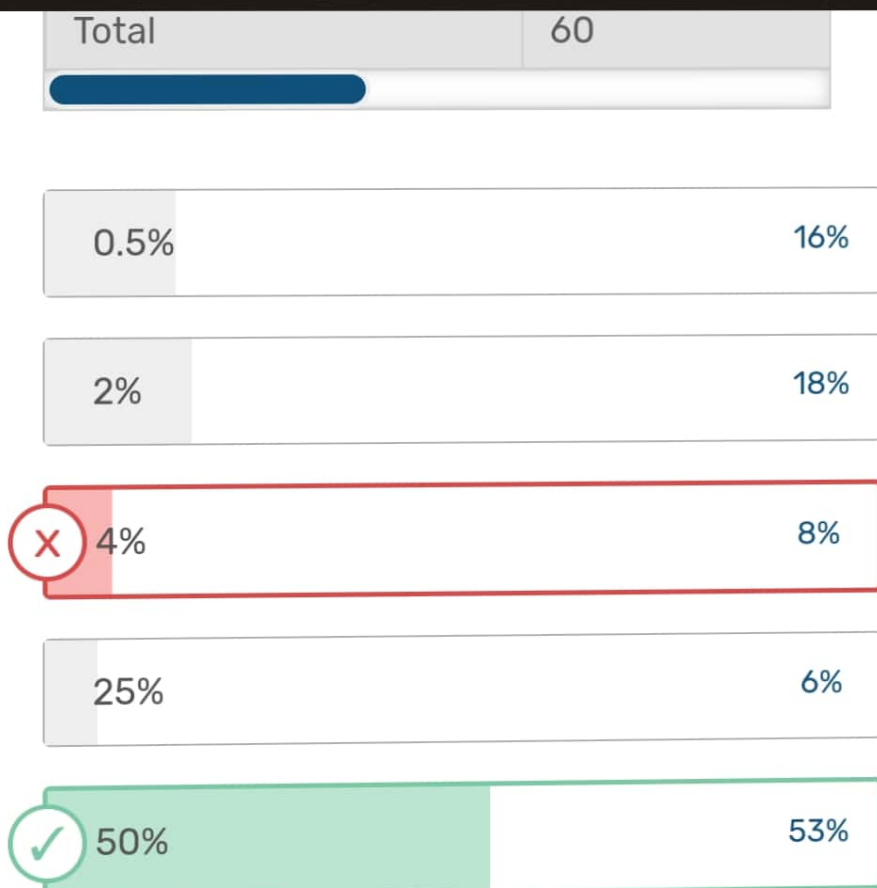
4%

25%

50%

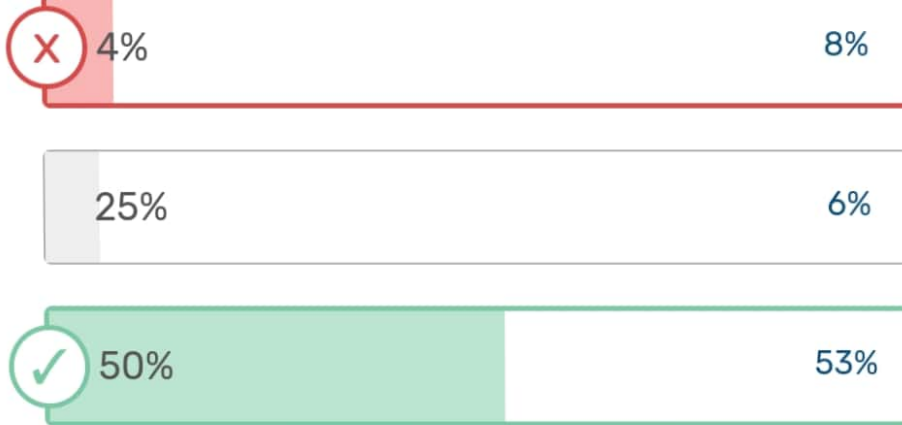


See Answer



## ANSWER

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the absolute risk of having an MI by 2%.
- The RR =  $ART/ARC = (20/1000)/(40/1000) = 0.5$  (50%)



## ANSWER

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the absolute risk of having an MI by 2%.
- The RR =  $ART/ARC = (20/1000)/(40/1000) = 0.02/0.04 = 0.5$  (50%)
  - therefore the risk of having an MI while taking Superstatin is 50% that of having an MI while not taking Superstatin.
- The RRR =  $1 - RR = 1 - 0.5 = 0.5$  (50%)
  - therefore Superstatin reduces the

## ANSWER

---

- The AR of having an MI in the treatment group (ART) =  $20/1000 = 0.02$  (2%)
  - therefore the risk of having an MI while taking Superstatin is 2%.
- The AR of having an MI in the control group (ARC) =  $40/1000 = 0.04$  (4%)
  - therefore the risk of having an MI while not taking Superstatin is 4%.
- The ARR =  $ARC - ART = (40/1000) - (20/1000) = 0.04 - 0.02 = 0.02$  (2%)
  - therefore Superstatin reduces the absolute risk of having an MI by 2%.
- The RR =  $ART/ARC = (20/1000)/(40/1000) = 0.02/0.04 = 0.5$  (50%)
  - therefore the risk of having an MI while taking Superstatin is 50% that of having an MI while not taking Superstatin.
- The RRR =  $1 - RR = 1 - 0.5 = 0.5$  (50%)
  - therefore Superstatin reduces the relative risk of having an MI by 50%.



Save

End Session



## QUESTION 119

The results from the study investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the positive predictive value calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

$$a/(a+c)$$

$$d/(c+d)$$

$$a/(a+b)$$

$$(a+c)/c$$



The results from the study investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the positive predictive value calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

$$a/(a+c)$$

$$d/(c+d)$$

$$a/(a+b)$$

$$(a+c)/a$$

investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the positive predictive value calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$$d/(b+d)$$

2%

$$a/(a+c)$$

15%

$$d/(c+d)$$

4%



$$a/(a+b)$$

78%



$$(a+c)/a$$

2%

## ANSWER

Positive predictive value (PPV) is the proportion of individuals with a positive test result who actually have the disease.

$$PPV = a/(a+b)$$

A new test is being developed to diagnose chlamydia. 1000 people aged 15 – 35 years attending a GUM clinic undergo the new test and the current gold standard nucleic acid amplification test (NAAT) to confirm the diagnosis. Of the 1000 people, 250 are diagnosed with chlamydia. Of the patients diagnosed with chlamydia, 240 test positive with the new diagnostic test and of the patients not diagnosed with chlamydia, 150 test positive with the new diagnostic test. What is the sensitivity of this test:

|               | Chlamydia Yes |
|---------------|---------------|
| Positive test | a= 240        |
| Negative test | c = 10        |
| Total         | 250           |

62%

75%

80%

96%

98%

|               |               |
|---------------|---------------|
|               | Chlamydia Yes |
| Positive test | a = 240       |
| Negative test | c = 10        |
| Total         | 250           |

62%

8%

75%

4%

X

80%

7%

✓

96%

67%

98%

13%

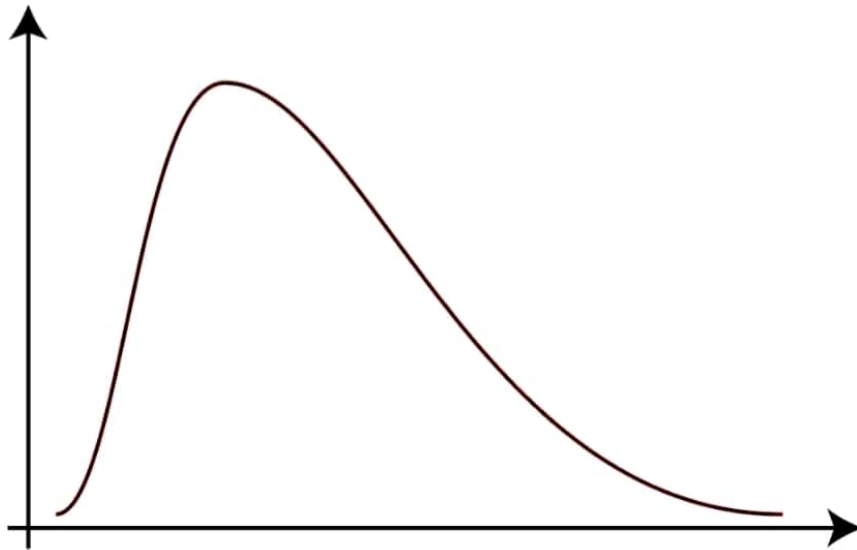
## ANSWER

- Sensitivity =  $a/(a+c) = 240/250 = 0.96 = 96\%$ 
  - This means that if the patient has chlamydia, there is a 96% chance of the test being positive. The test will only have a 4% false negative rate.





Regarding the following data distribution, which of the following statements is CORRECT:



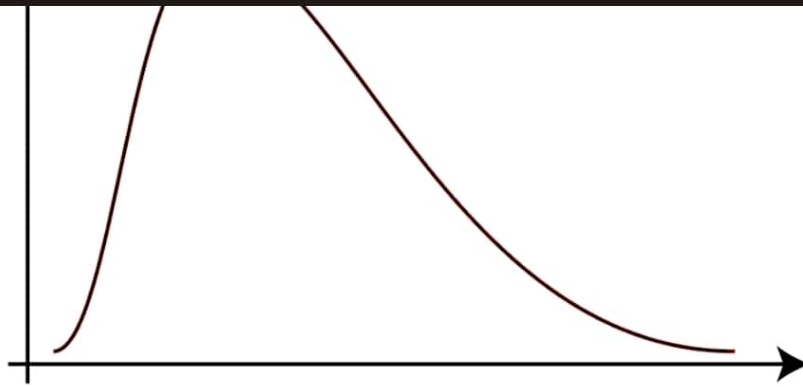
The distribution is negatively skewed.

The distribution is uniform.

The mean = median.

The mode is greater than the median.

The mean is greater than the mode.



The distribution is negatively skewed. 24%



The distribution is uniform.

2%

The mean = median.

3%

The mode is greater than the median. 18%



The mean is greater than the mode.

53%

## ANSWER

This unimodal distribution is positively skewed. In a positive skew, the right tail is longer and the mass of distribution is concentrated on the left. The mean  $>$  median  $>$  mode.



Save

End Session



Which of the following best describes the sensitivity of a diagnostic test:

The proportion of individuals with a positive test result who have the disease

The ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease

The proportion of patients with the disease who are correctly identified by the test as having the disease

The proportion of individuals with a negative test result who do not have the disease

The proportion of patients without the disease who are correctly identified by the test as not having the disease

the sensitivity of a diagnostic test.

The proportion of individuals with a positive test result who have the disease 19%

The ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease 5%



The proportion of patients with the disease who are correctly identified by the test as having the disease 66%



The proportion of individuals with a negative test result who do not have the disease 4%

The proportion of patients without the disease who are correctly identified by the test as not having the disease 6%

## ANSWER

Sensitivity is the proportion of patients with the disease (true positives) who are correctly identified by the test as having the disease = true positive rate.



The results from the study investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the positive likelihood ratio calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

$\text{sensitivity}/(\text{specificity} + 1)$

$\text{sensitivity}/(1 - \text{specificity})$

$\text{specificity}/(1 - \text{sensitivity})$

$\text{specificity}/(\text{sensitivity} + 1)$

$(1 - \text{sensitivity})/\text{specificity}$



The results from the study investigating the accuracy of a new diagnostic test can be displayed in the following format. How is the positive likelihood ratio calculated:

|               | Those with disease |
|---------------|--------------------|
| Test positive | a                  |
| Test negative | c                  |
| Total         | a+c                |

- ☐  $\text{sensitivity}/(\text{specificity} + 1)$  5%
- ☒  $\text{sensitivity}/(1 - \text{specificity})$  78%
- ☐  $\text{specificity}/(1 - \text{sensitivity})$  9%
- ☐  $\text{specificity}/(\text{sensitivity} + 1)$  4%
- ☐  $(1 - \text{sensitivity})/\text{specificity}$  5%

## ANSWER

The likelihood ratio for a positive test result (LR+) is the ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease.

|               |     |
|---------------|-----|
| Test positive | a   |
| Test negative | c   |
| Total         | a+c |

- ☒ sensitivity/(specificity + 1) 5%
- ☒ sensitivity/(1 - specificity) 78%
- ☐ specificity/(1 - sensitivity) 9%
- ☐ specificity/(sensitivity + 1) 4%
- ☐ (1 - sensitivity)/specificity 5%

## ANSWER

The likelihood ratio for a positive test result (LR+) is the ratio of the chance of a positive result if the patient has the disease to the chance of a positive result if they do not have the disease.

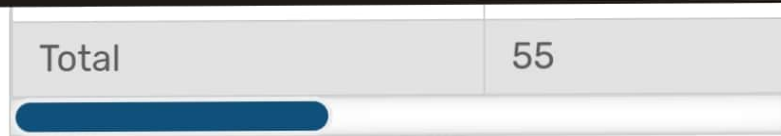
For ruling in a diagnosis,  $LR+ = \text{sensitivity} / (1 - \text{specificity})$





A cohort study is performed, where researchers follow 100 patients with a high alcohol intake and 100 non-drinkers for 20 years to look for the development of liver disease. 50 of the patients with a high alcohol intake and 5 of the non-drinkers develop liver disease. What is the risk ratio for developing liver disease in this cohort:

|                     | Liver disease p |
|---------------------|-----------------|
| High alcohol intake | 50              |
| Non-drinker         | 5               |
| Total               | 55              |



## ANSWER

The risk ratio for developing liver disease in patients with a high alcohol intake compared to non-drinkers = Risk of liver disease in patients with high alcohol intake/risk of liver disease in non-drinkers

Risk of liver disease in patients with high alcohol intake =  $50/100 = 0.5 = 50\%$

Risk of liver disease in non-drinkers =  $5/100 = 0.05 = 5\%$

Risk Ratio =  $(50/100)/(5/100) = 10$ .

Therefore patients with a high alcohol intake are

15

10%

X

20

10%

25

9%

## ANSWER

The risk ratio for developing liver disease in patients with a high alcohol intake compared to non-drinkers = Risk of liver disease in patients with high alcohol intake/risk of liver disease in non-drinkers

Risk of liver disease in patients with high alcohol intake =  $50/100 = 0.5 = 50\%$

Risk of liver disease in non-drinkers =  $5/100 = 0.05 = 5\%$

Risk Ratio =  $(50/100)/(5/100) = 10$ .

Therefore patients with a high alcohol intake are 10 times more likely to develop liver disease than those who do not drink based on this cohort.



Save

End Session



In hypothesis testing, what does alpha denote:

The chance of making a type I error

The chance of making a type II error

The chance of detecting a statistically significant difference where one is present

The chance of correctly accepting the null hypothesis when it is true

The chance of correctly rejecting the null hypothesis when it is false



See Answer

Save

End Session



In hypothesis testing, what does alpha denote:

- ☒ The chance of making a type I error 66%
- ☐ The chance of making a type II error 5%
- ☐ The chance of detecting a statistically significant difference where one is present 16%
- ☐ The chance of correctly accepting the null hypothesis when it is true 8%
- ☐ The chance of correctly rejecting the null hypothesis when it is false 5%

## ANSWER

The probability of making a type I error is denoted by alpha ( $\alpha$ ), the predetermined significance level of the test (conventionally  $\alpha = 0.05$ ).





Box and whisker plots are typically used to display:

The interquartile range

The existence of publication bias in meta-analysis

The survival of a sample cohort

The strength of evidence of constituent trials in a meta-analysis

Confidence intervals



See Answer

Save

End Session

SCORE

14%



Box and whisker plots are typically used to display:

- ☒ The interquartile range 72%
- ☐ The existence of publication bias in meta-analysis 6%
- ☒ The survival of a sample cohort 4%
- ☐ The strength of evidence of constituent trials in a meta-analysis 12%
- ☐ Confidence intervals 6%

## ANSWER

Box plots or box and whisker plots are used to display the median and interquartile range.



Save

End Session





Regarding hypothesis testing, which of the following statements is CORRECT:

The null hypothesis assumes that any observed difference between two groups is due to chance.

The p value gives the probability that an observed difference between two groups has happened due to a true difference.

$P = 0.5$  means that the probability that an observed difference has happened by chance is 5%.

The closer the p value is to 0, the more likely it is that a difference has happened by chance.

Conventionally significance levels are set at 0.1.



✓ The null hypothesis assumes that any observed difference between two groups is due to chance. 44%

The p value gives the probability that an observed difference between two groups has happened due to a true difference. 24%

✗  $P = 0.5$  means that the probability that an observed difference has happened by chance is 5%. 17%


The closer the p value is to 0, the more likely it is that a difference has happened by chance. 12%

Conventionally significance levels are set at 0.1. 3%

## ANSWER

The null hypothesis assumes that there is no true difference between two groups of data and that any difference observed in the results of two groups of data is due to chance. The p value gives the probability of observing a difference between the two groups if the null hypothesis is

the p value gives the probability that an observed difference between two groups has happened due to a true difference.

  $P = 0.5$  means that the probability that an observed difference has happened by chance is 5%.

17%

The closer the p value is to 0, the more likely it is that a difference has happened by chance.

12%

Conventionally significance levels are set at 0.1.

3%

## ANSWER

The null hypothesis assumes that there is no true difference between two groups of data and that any difference observed in the results of two groups of data is due to chance. The p value gives the probability of observing a difference between the two groups if the null hypothesis is true. In other words, the probability that any observed difference has happened by chance and not by any true difference.





A randomised controlled trial (RCT) is carried out to study the effects of a new anticoagulant, Anticlot, on stroke incidence in patients with atrial fibrillation (AF). 500 patients with AF are treated with Anticlot and 500 receive usual therapy (control group). One year later the number of patients who have had a stroke is recorded and shown below. What is the relative risk reduction of treatment in preventing stroke:

|          | Stroke Yes |
|----------|------------|
| Anticlot | 30         |
| Control  | 50         |
| Total    | 80         |

new anticoagulant, Anticlot, on stroke incidence in patients with atrial fibrillation (AF). 500 patients with AF are treated with Anticlot and 500 receive usual therapy (control group). One year later the number of patients who have had a stroke is recorded and shown below. What is the relative risk reduction of treatment in preventing stroke:

|          | Stroke Yes |
|----------|------------|
| Anticlot | 30         |
| Control  | 50         |
| Total    | 80         |

4%

6%

10%

40%

60%





10%

10%



40%

46%

60%

17%

## ANSWER

- The AR of having a stroke in the treatment group (ART) =  $30/500 = 0.06$  (6%)
  - therefore the risk of having a stroke while taking Anticlot is 6%.
- The AR of having a stroke in the control group (ARC) =  $50/500 = 0.1$  (10%)
  - therefore the risk of having a stroke while not taking Anticlot is 10%.
- The ARR =  $ARC - ART = (50/500) - (30/500) = 0.10 - 0.06 = 0.04$  (4%)
  - therefore Anticlot reduces the absolute risk of having a stroke by 4%.
- The RR =  $ART/ARC = (30/500)/(50/500) = 0.06/0.1 = 0.6$  (60%)
  - therefore the risk of having a stroke while taking Anticlot is 60% that of having a stroke while not taking Anticlot.
- The RRR =  $1 - RR = 1 - 0.6 = 0.4$  (40%)

## ANSWER

---

- The AR of having a stroke in the treatment group (ART) =  $30/500 = 0.06$  (6%)
  - therefore the risk of having a stroke while taking Anticlot is 6%.
- The AR of having a stroke in the control group (ARC) =  $50/500 = 0.1$  (10%)
  - therefore the risk of having a stroke while not taking Anticlot is 10%.
- The ARR =  $ARC - ART = (50/500) - (30/500) = 0.10 - 0.06 = 0.04$  (4%)
  - therefore Anticlot reduces the absolute risk of having a stroke by 4%.
- The RR =  $ART/ARC = (30/500)/(50/500) = 0.06/0.1 = 0.6$  (60%)
  - therefore the risk of having a stroke while taking Anticlot is 60% that of having a stroke while not taking Anticlot.
- The RRR =  $1 - RR = 1 - 0.6 = 0.4$  (40%)
  - therefore Anticlot reduces the relative risk of having a stroke by 40%.